```
AGCACACACA GGCGCGTCCT TCCTCTTTTT GGGAGGATCC GCTATGCTCA GAGCCATGCG        120

CACCATTCCC CCTATCACCG GCATTATTGC CTCAGCAGGA TGCGCCATCG GTCCAGTCTT        180

TTGCTTCGAT ACCCTGCTAC CTACCCGCTC CCGCCCTGGC GGATCCCGCC TGCGCCCCCC        240

ATCGCAGGAA ATTGCCCGCT TGCGCAACGC ACTTTCATAT GCGCGCGCCT CGCTGCAGAA        300

CCTGCTCGAT TCAGTGCGCG CAAAAGCATC CGGAGACGAG CCTGCACCCG AGTATGCCGT        360

GCTCAGTGGC CAAGCAGAAA TGCTGGCCGA CGCGGCCTTC ATAGCTACCG TAGAGGAAAC        420

GCTGCGCTCT TGTTCCTGCG ATGCAGAAAC TGCTTTGCGC AAGGCAATTA CCCACGTGAC        480

AGATGCCCTC TCTGCTACCT CAGACGAGTA CCTGCGTGCC CGGGCAGCCG ATATCCGAGA        540

CGCGTTTAGG GTGTCTTCGA CGCACTTGCG CATGACACCA CACCCACCGC AGnAAGCTCT        600

TTGCCAACAC AAGGGATTGG AAATAGCACC CCACACTCCC CCTGGGAGCC TGACTTTAGC        660

GCCGTTCCCC CAGGATCCAT CGTGGTTGCC GCTCACGTAC AACCTGCGCA CGCACTGCGC        720

CTGCACGAGG CAAATATCGC TGGTTTGGTA ACCGAAGTGG sCAGCGTAAC AAGCCATGTC        780

GCCATCATGG CGCGCGCGTG GAGTCTTCCC CTGCTCGTCA GTGCACAGGG ATGTAAAGAC        840

GTTGCACAGT ACGTGCTCCG TGTGCGGCAA ACTGCTCGTG CCACCGATGA GGCGCTGCGC        900

GCACTCCTCG ATGCTGAAAs AGTGGGGGAA AAACTGACGC TCTAGGAACC CTCACCGTAA        960

ATCCCGACGT GCGCGCGCTG CgCACrCGCA TGCCTcACCC TTTCCTCACC GTCAAACACA       1020

CCAGTACAGC TGAACAGAGT CCCCCGGCCG CCTGTGTGCT AAACGCACCG CTGCGCACTT       1080

ACTCAAGTGA CGGTATCCGT TTTGAAGTCG GGGCAAATAT CGTTATGCCC CAGGAAGCGT       1140

GTGCAGCTGC TGCGCTCGGA GCAGCAGGCA TCGGACTGTT CCGTTCGGAG TTCTTGCTAT       1200

TCGGATCCGA CCGCTTCCCA GATGAAGAGA CGCAgTGCTC TGCCTACACG CGCGCGCTGC       1260

AGGCAATGAG AGGACTCCCC GTCGTGCTTC GAACGTTTGA CCTTGGTGCA GACAAACTGG       1320

TGCCAGACCC TGCGCGAATG TGCGCACTCT CGGACGCTGC TGAACCGTGT GCACACACCG       1380

CTTCGGAGCG CAATCCTCTT TTAGGGTTAC GAGGCATCCG CTACTGCCTC GCACATCCTG       1440

AGCTCCTGAA AGTGCAGCTT CGTGCAATGt CCGCGCCGGA rCkTGCGCAA CATGTGCAGA       1500

AGGGnACTGC GCATTCTCAT CCCCATGGTT TCACGGGTGG AAGAAATTCA CGCCGTCGCC       1560

GACCTCATCT CTGAGGTAGC CGAcGAgTGT GCCCGCGCGC ACGTGAGTAC ACCCGATCGG       1620

GTAGCACTCG GCATTATGAT CGAAACGCCC GCTTCGGCAC TGATGGCAGC AGAttCGCTC       1680

CCCACGTGGA TTTTTTTTCC ATAGGGACGA ACGACTTAAC CCAGTACGTG TTCGCCGCCG       1740

ATCGAGAAAA CGAACAGGTC AGCAGCTATG CCGATTACTT CCACCCGGCA CTCCTCCGTC       1800
```

```
TTATCCAGCA CGTAATACAT GCGCACAGAC ATCTGCGGCA ACGTCCCGGT ATTTCTTTTG    1860

GAGAACAGGG AATCGGACGC GTGGTCATGT GCGGCGCCAT GGCTGAAGAT GAAAtGCGCT    1920

CTTTCTTCTG GCGGGGCTCG GCCTGCGAGC GTTGAGTGTG CCTTCTTCAC GCATCGAGAC    1980

GCTGCACACG TTCTTATCAC GCATTTCAGT CTCTGATGCA GAGCACTGTG CACGTGCAGC    2040

CGTGCAGCTT TCAGATGCGC AGTCAGTCCG CACACTCATC GAAGAACATC TGCGCACCGC    2100

AGGTATTACG CTTGAGAAAG ACGAGGAAGA ACCCTCACCC CCTCGATCCC CATAGCGGAG    2160

GAGGCCTCAG GCGTTTTCTC CATACACAGA AAGGAAAAGG CAATGGAAAT CGAAGAATTT    2220

GGTCCACAAA TCACCGCCCT CGAGGCGCGC GTGCAGGAAG TATGGGGGAG TCTTTGACGT    2280

TGCCGCATAC GAGGCGCGCA TAGCAACGCT TGAgGCTGCT GCAGCAGCGC CTGACTTTTG    2340

GAGCGAACGC GCGCGTGCCG AAGCGCTGTT AGCGGAACTG AAAAAACTAC GCGCAACGCT    2400

TGAGCCGTGG CGTTGCGcTG CGCCGTGAGA GCGCAGATCT GCGCGCGTTG TACGAGCTTG    2460

CCCGCGAGGC GCAAGACGCA TCGCTGGAGC CAGAACTTTC CTCCCTTTTT TCAGACATTT    2520

CTGCTCGTTT CGAAGAGGCA TCGCTTACCC GTCTCCTGCA CGAAGAGGTA GACCGCCTCG    2580

ACGCGTTTGT TACCATCCAC TCCGGCGCAG GAGGAGTGGA GGCCTGCGAC TGGGCACAGA    2640

TGCTCATGCG CATGTACACG CGCTGGGCAG AGCGGCGCAG CTTTTGCGTA CACATAGTTG    2700

ACTTACTTGA GTCAGAAGGG GGAGTAAAAT CGGTGACGTT AAAAATTTGC GGGTCACACG    2760

CCTTTGGTTT TCTCAAGGGA GAAACGGGGG TACACCGGCT CGTGCGCATC AGTCCGTTTG    2820

ACTCTGCCGC GCGCAGACAT ACCTCTTTTA CCTCCACCTA CGTCTTCCCC GTATTAGACG    2880

ATCACGTTGA GGTGCACATA CGGAGCGAAG ACATGCGGGT AGATACCTAC CGCTCAGGGG    2940

gAGCAGGCGG TCAACATGTC AATAAAACGG ACTCTGCCGT GCGCATCACG CATCTGCCTA    3000

CAGGgATAGT AGTCACCTGC CAGAACGAGC GCACCAAATC AGCAACCGTG CAAgGCGCTG    3060

AGCTTGTTAC GCGCCCGCCT GTACGCCTAT GAACGGCAAA AAAAACAGCA GGAACATCAA    3120

CGGTTTGCTT CTGAAAAGAA GGATATTTCG TGGGGAAATC AGATTCGCTC GTACGTCTTT    3180

CATCCCTACA CCATGGTTAA AGATCACCGC AGCAAGTGCG AAACGGGGAA TATTCACGCA    3240

TCATGGACGG AGCGTTAGAA CCGTTCATCC GTTCCTACTT GGAGTTTCTG TGTACCAGTA    3300

CCCAGTGTGT AGAACCACAG TGAACGGGAG TTACGCGCAA TCATTTGCAG CACTGCTTTT    3360

CTTTCCCCAA ATCGCGGTCG gTTTAGTGCA AnGGCACCGG cGCCGTCCTT TGACTCTTTC    3420

CTGTCCGCGC GTCAGTCcAC CCTCCTGCCT CTTCCTTTCT AGCATCACCT GCAGCGCCGA    3480

CACACCCTCT TTTCGCAACC GCCCGTACGA CTGCTGCGCA cTGCTGtCTA CGCCTCCTGC    3540
```

```
nCCCGCCCGT CCCCGCCCGC GCCCCGATCA CGTAATCCCC AAGGAAAAGT GGCGCCTTGC    3600

CGTTGCAGAC TTTACCTTTC ACGGTATTCC AAAGATTTTT CAGCGCTACG TGCGTCCTGC    3660

GCGGGAGctA CTCTTTATTG AACTAAAAAA ATTACCCCTC CGTCATTTTC TTTCTGAAGC    3720

TGAACAGCGC GAGcGCGCCG CCTTGCCCCA CGAAGAAGCC TACCACGCCC GGCTCAAAGA    3780

ACGTGCACAT TTACAGCGsG CGCGTGATTT TGTTTCCTTG CACCCTGTCA GCGATCACGC    3840

GCGCCGTCTG CGTACGGCAG CATTTGAAAA GCAAATCAAA GAGAAGGAGC AAGAAATCGA    3900

GCGTGCCCGT GTGGAAGTGC GCACgCACGC GCGCGGTTTT TCCGTCCCTG GCTCCAGGCA    3960

GAGGTGCTCG TCTTAGGTGC GCAAAACGAA CCGCATGCAC TGCCTGAGCG CTTTCACCTT    4020

GCCACCCATT TACGGCAAAA AAAACTTTCT GCACTGGTTA CGGGAAAACT CGTAGACGTC    4080

GCCGGTTACG TGCGCATATC TCTCTATCTT TCTACAGGGC TAGAAGCAGA ACCCACGCGG    4140

GAATTCACGC TCGCAGGTCC CTACCGAGAA CTGCCGCGTC TTATGCACAC GCTGTCTGCA    4200

CAATTGCGCA GTGCCATTGA AAACGCACAA CCGGTGCGCA TTGTGTTTGA CGTACATCCT    4260

CCGCATGCAC GTCTTTCGTT TCAGGGCGTG CCGGTAGAAG ACCTTTCCAA ACCTCTTATC    4320

TCATACCCGG GCCGCTACGT GGTGGACGTG TCTGCTGCAG GATACTTTTC TGCCACAAAG    4380

GAAATATACA TTGAAAACCG ACCTGCCTTT TCACTACGGG TGCGTTTAGT TGCCCGTCCA    4440

CAACATCGTG TGCGCGTGCA GCTTACTGAC AACAGCGCAG cACCTATCTT TTCTGGCGCA    4500

CGCTCAGTGG GAGTCACTCC CTTCAGCACC GTGGTTACTG ACTTGCGCGA AATTTTCACC    4560

GTCGGACCGG CAGGCGCGCG TTCGTTTGCC TTCATTGAAC GCGGCACATT TCCTAACTCT    4620

CAGCCGAGCA CGCTCGTGTT GCCTGCGCCT AACCCAAACG CAACACAGGA TCTTGCGTAC    4680

AAAAGGGACG TAGCATACTG GTCTTTTGGA GCCCTCTGCA TTGCCGTTCC CATCGCGCTC    4740

ATTCTCGGCT CCACGCTTGC AGACACGCAT CAGGCGCTAG AACGCGCAAA AGCTGCAAgC    4800

GCGgCAACCT CCTCCCCCTC CTGCACCGGC CGGCACGGGC GCATTAGAAC GTAAAAGCCA    4860

GCACCTGCTC ATCGGCACGG GGGTAGCAGT AGGAGTGGCG GTTATCCTGA GCATTAATTT    4920

CATCGTGCcA CTGCGCGCTA TTTGAACGCG GTGATGCACA ACGCGCCACA GGCAGTACGT    4980

CCCCGCGCGG ACAAAGACAT ACAAACATTA ACGCACCGCG ACGAGGCAGA AGAAGATCAG    5040

GAAGAAGATT CCTAAAGGAG CGTGAAGGTG GGTTTAGGAA ACCTAGCACA GAAAATACGA    5100

CGCCTGCTCG GTGGACAGGC GCCTCTGGAC GAAACGTTTT TTAGCGCGCT TGAAGAGCTG    5160

CTCATCGAAG GCGACCTGAG TCTTTCGACG GCAGAGAGCT TTTGCACACA GCTTCGAAAC    5220

GCCGCGCGCA CACGTTCTGT ACATACGGAA GACGCAtGCG CACGCTCTTT GCGGAAATTA    5280
```

```
TGGAATCGTG CGTACGCGTT ACCCATCTTG CACCAAATCC GAACCAGTGC TCACTGTATC    5340

TCCTACTTGG GGTTAACGGG AGCGGGAAGA CCACTTCTGC TGCAAAGTTG CAGCGTACTA    5400

TCAGACCCAG AAGGTGCATC CGATACTGTT TGCCGCCGCA GATACGTTCC GCGCAgCAGC    5460

GGCAGAACAA CTCGCACACC ACGGTGCACA GCTAGGCGTG CGCGTCATTG CGCACCCGGG    5520

GGGAAAAGAT CCTGCTGCaG TGGTATTTGA CGCAGGAGAA GCCTTGCgcG cGCAAAAGCG    5580

sGGTCTTTTA CTCGTTGACA CCGCAGGGCG ACTGCACAAT AAGACGCACC TCATAGCGGA    5640

GCTGCAAAAG ATCGACCGTA TTGCGCAGAC AAAGGTGAGC GCAGATGCAT ACCGCAAGAT    5700

ATTGGTATTA GATGCCACCA CCGGTCAAAA TGCATTTCGT CAAGCGCAAA CtTTCACGAA    5760

GCTATTGGCG TGGATGCACT GCTCCTTGCA AAATGCGACA CACGCGCACG AGGGGGAGCA    5820

GTTTTTTCCA TCATGCAAGA GTTAGGTATT CCATTAGCCT TTTTAGGGTG GGGGGAGCGC    5880

TATACAGACT TGGTTGAAGC GAACGCGCGC GAGTTTGTTT CCTCGTTCCT GCACGGAGAA    5940

CGATGATTCG ACCCCGGTAT GGCTGGATGT ACAGCAGCGG GATTGCAGTG CACCTGTGTG    6000

CgGCCgTGTG CGCGCACAGT GCTGTTCCTG CCGCGTGGAC CTTTGCAGAA CAGACACAGG    6060

CGCAAAAAAC AGACACTCCG CTTGATTCCT CCaGtACGCA tGaCCTCCCC TGAGGAAGCA    6120

CCCAATGAAG CAGATCCGTT TGAGAAGGAA CTGGaACACG CGTTCGAAAG AGCGCACGTC    6180

AGCACAGGCG GTGCAGATTC CTCATCACAC GCCGATTTTG TACACATGGA AGAGGCAGGA    6240

CGTGCCCACG CGTCCGCCAA TCGCTGGTAT CACGAAACGT TTGACTCGCG TCAGCGTCCA    6300

TCCTCTGCAG TTCTGTACGA AGGGGCACAG CTACTGCATA CCGTTCACTG GCACTATGTC    6360

gGGGACCGGC TGTTTCCCTG TGAAAAAATA ATTACCACAC CACACACACG TATnCCGCGC    6420

GCGCTATAAT TTTTCCGGAA AGATCGTCGC GTACGAAATG CACACGCGCG GGGTACTGGT    6480

ATACGCACGC ACCTATCGGT ATGnAnCGCA CGCGCGTATA TGTGAAAAGG AAGAAACAAC    6540

TGCTCGAGGG AATGAACGCA TTACGTATGA                                    6570
```

(2) INFORMATION FOR SEQ ID NO: 42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19483 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

```
TTTTTGCGCG CGTTCTAGCA CCCGAGTnAA TAGTGTTTTT TGAAAAATGG AGGnTCGCGT    60
```

```
CTACCCAGTT GTAAAAAGAG tGTTCGCGCG CGTCCgTGCT CTCCACACGA AcGGAcTCCG     120

TCCACTCACG AAAGATATCG TGTCCGAAGT ACAATACGAG CATCATAACG TAAACTGCAC     180

AGGGTGTTGT GTACgtsTGt GCGCACATGT AGCACCTTTC CCATACGGAG ACACAGGGAG     240

AAAAGTTCCA AACGGGATGT GCACGCGTAA AAAAGTGAAG ATGCGCACAG CAATAAAAAT     300

AGGCGGATGC CTATGTGGTA TCCTGCATGT ATGTGCTGGT TGATTTGTGT AGCGCTAACA     360

ACACCGGTGC TGAAAAATGT GTGCAGGATC TTCAGAGAAA AGAAAAAAGC GAGGTAGATA     420

CTAACTACAC GTATGTGTGC AGCCACGCAT TTCAGAGAAC GTGTGACAAT GAGCGTTAAT     480

GCGAGCAGCA CAAGTGTAAG CGACGCgTGC ATACACCAAC TCTGTGCATT ACAGCACAGG     540

AGAAGGAGAG CGAAgcTTGC GACTTTCGCT ACGGGAGGAG CACGGTGAAG CAGCGATTGC     600

CGGCGTTCAT AAAGAGAGAA AAACATACAT TTCCTTGACT CCTCCTAGCA GGTGGAAGTA     660

CTGCATGCAT GTCGCTACCT AGGTCCAAAG GAGATCTGAA AGAGTGAGTG GACAGTGCAA     720

AGGATCTCTT AGTCCGTAGC AAGAGAAAAT TTTACTGTCG AGTGCGTCCT GTGGCGTGCC     780

ATCGTAGGAA ATGACCCCCT TAGAAAGGAT GCATAGACGC GTCGCAGCAG CGAGTATTTT     840

TTCAACCTCA TGGGTGATGA TAACAAGCGT TTTACCTGCG TGTTTGAGGC TTATGATGAG     900

CTGCACAACC TGACGAACGC TGGGGTAATC TAAGTTTGCA AACGGCTCGT CAAGAATGAC     960

TACCTTTGCA TCCAAGGCGA GTACGCCGGn CAACGGTTAG GCGTCTTTTT TCTCCACCTG    1020

AAAGCGCTCG GGCGTAATGG TCACGCCGGT CAAGCAGTGA CACGGcTGCA AGTGCGCTGT    1080

TGGTACGTGC GTCAATTTCT GCGCGGGAAT ATCCCCACTG CAGAGGACCG AAGgCGCAGT    1140

CCTCAAATAC CGTTTCGCCT AGGATCTGGG TGTCTGCATT TTGAAACGCC AGACCGACAG    1200

TAGTCCCGCG TGCCATATAC ACACGGCCGG AGGAcGGCGG TTCAAGTCCT GCAAGaTGTT    1260

CATGAGCACA GTTTTACCCG AGCCATTTGC ACCTGCGAGG ACGACACAGT CCCCAGGAAA    1320

CACCCTAAAC GAAACGGAGT GTnAATACTT CACAGTCGCG CTCAAAAGAC TTACTTACAT    1380

TGACCAGTTC AAGCAGCGGT CCTGCGCACG ACTCCACAGC CGTGTCTGCC GCGACATCTG    1440

CGCTCATGCG TCCACAGAAG ATTCACCGTG CCCTGTGGCG CGTACACACC CGCGACACGC    1500

GAATACTATG GCATCAACCA TGACGGTACA AATGGCGGCG AACTGTAGGG GCAAGATGAT    1560

GGGCGAGTAG GACAACAAGG GCGATTTTCA GGGTGTCGGC AAGAAAAAAG GGAAGGAAGA    1620

ATCCTAGCAT GAGCTCCCCG GTCtTGAGGC CAAGCACGTA AcCGAGAACC GgCAGACCGA    1680

TGGAGTAAAT CGAAAGAAAA CCAACGAGCG TTGCGACGnT AAGTCTGATC CAAAGAAGGA    1740

GTGCGCGTTC CACCACCGCG TGGTGGAGTG CAAGACAGGC GGTGGTGCTG CGCGATTGCc    1800
```

```
CAcGAGCGTA GCAGCGAGTA TGTATCCAAG GAGGAATCCT CCCGTAGGGc AAAAAGCGCG    1860

GTGTATCCGC CCCGACCTCC TGAAAAAACC GGCAGACCAA GGAGTCCTGC CCCGAGGAAG    1920

CTGAGAACGG CGAGTGCACC GTCTCGCGGT CCCAACAATA AACCGGTGAG AACGGCCGCT    1980

GCATTCTGCA GTACAAGCGG AACAGGCTTG AGAGGAATGC TAACGAGCGC ACTCGAGCTA    2040

ATGAGTGCGG CAAAAAGCGC AACAAAAGCC AAAGACTTAC TACGGTGCAT GGTACAGTAC    2100

TCCTCCGAAG GTTCGATGCG CAGTGTGACA CGGAAGGAGG AATCTTTCAA TATCTTGGGT    2160

GGTGCCACAG GTATAGTTTT TAACAGACTT ACCCGAACGG CTGCCAGGTG CGTACGCGTC    2220

GGTTCGTTCC CACTGTGCGC GGGCAGAGCT CGTGTAGTGT CTATTGACAG ATGCAAGGAT    2280

CGGGTACCGT CATGTACGCA gTTTATGGTA GGCTCAGTCC TATGCACGCT GGGGACAGAG    2340

AGAGTATCGC ACGCTTCGTG cGTGTGGTGC GCGATTGTCT GGATTTGTTT CGCACCGAGG    2400

GTATTGGGCC CCGTCCTAGG AATGATTCGG TAATTTTACC GAATGCTGCG TGTTCACCGC    2460

GTAATCATGC AGGAAAGCGT GCGCAGAGCA CTGCCGATGC GTGTGTGAGA AGCAGTGACG    2520

GGTCTGTATA CACGGACGAA ACCTTGCGCG AGGAAATTTT TGCATGCCGT GCGTGTGAAT    2580

TGTATCAACG GCGTACACAT GCGGTGGTGG GAGAGGGTGT TGCAGACGCA GACGTGCTCG    2640

TCGTTGGGGA GGCCCCTGGA GCGGAAGAAG ATCGAAGCGG TCGTCCGTTC GTAGGACGGT    2700

CAGGTAAATT GCTGGACGCA ATGCTTGCGG CGATTGGACT TTCGCGTCAg cAAAATTGTT    2760

ATATCACCAA TGTGGTTAAG TGCCGGCCGC CAAGGAACCG CACACCAACA CCCCACGAGA    2820

CTGCCTGTTG TGCACGGTTC CTCCATGCGC ATCTTACGCT GCATCGCCCG TGTGCTATTT    2880

TGGTGCTCGG CCGcTGCGCC GCACAGCACA TGCTCCAAAC AACCGATGGT ATTGGCAAGT    2940

TGCGCGGgCG CTTTTTTACC TAtCAGGGgA TtCCCCTTCT GGcTAcGTAC CATCCGAGTG    3000

CGTTGTTACG GGATGAAGCG CTGAAACGTC CGGCGTGGGA GGATCTCAAA ACGTTCGTG    3060

CACGGTTGCT GCAGTTGAAG CAGGACGCAC ACATGCCAAT ATAAAATCAT GGCGCCGTGG    3120

CTTGAGCTTG TTTTTGAcgT TCCACTGGAT AAAAGCTTTA CGTACCGTGC GTGTGCTGCC    3180

CACGCGGGTG AgGCACTCGT GGGTAGACGG GTTCTTGCTC CCTTTGGGGC GCGTACACTC    3240

ATTGGATTTG TGATAAGTGA ATCACATTCT TCGCCTGCTG ATTGCGGTGG TGCAGTTGGC    3300

ACGTTCAAGG AGATCATCCG CGTCATTGAC AGGGAAGCGC TTTTTGACCA AACGCATCTT    3360

GCGTGTGCGC GTTGGATGGC GCATTTCTAC CTGTGTGCCT TAGGTCAGGC GCTGTGTGCG    3420

GTGGTTCCGT CTCGGAAACG AGAACGGACA TTGTCTTCTT TTGCTTCTTG TGCGGGTGTT    3480

CGGCGCACTG ACACCTATGC GCTTTCGGGC GAACAGCGCA AGGCGATTGA TGCGATTACC    3540
```

```
GCGAGCACCG GTGCGCGCAg TTTTTATGTG CACGGGGTGA CAGGGTCGGG GAAGACGGAA    3600

GTGTTCTTGC GCGCACCGAG GCAGTCCTTG CGCGTGGCAA GTCGGTTATC TATCTTGTTC    3660

CTGAGATAGC GCTCACTCAC CAGGTGCTCC AGGAGGTATA TGTGCGCTTT GGCAGTCAGG    3720

CGGCGGTGTT GCACTCAGCG CTCAGTGGCA GTCAGCGCCT AGGTGAGTGG CGGCGCATAC    3780

AGTGCATGCG TCACTGTGTA GTGATTGGAG CTCGGAGTGC AATTTTTGCT CCGTTGAAGC    3840

GGCTGGGCCT TGTGATAATG GATGAAGAAC ATGACAGTTC GTATAAGTCT GCGCATGTGC    3900

CGCGCTATCA TGCGCGgCAG GTAGCGATGT ATCGCTGTGC GGACGCGAAC TGTCCGTTTG    3960

TCATGGGGTC TGCAACACCG TCTGTGGAGG CCTGGTACGC GATGCTGCGG GGGGCGGTGC    4020

GTCGTTTACC ATTGACTGCG CGTGTTGCGG GGGGGcTCCG CCGCGTGTTG AGGTGGTGGA    4080

CGTGTCAAAA GAGGCCCTGT TGCTCTCTAC CCGTCTGGTG GATGAAATAC GCAAGACGAA    4140

GGAGGCAGGA TATCAATCGA TGCTCTTTTT GAATCGTCGA GGATTTTCCT ATTCGTTTCA    4200

GTGTCGCAGC TGTGGATACA CGCTGTGTTG CACGCAgTGg CAGTTCCCTT GACGTGGCAC    4260

AAACGTGTGG GGGCAATGCA ATGTCATTAC TGTGGCAGGC AAGAGGCGCC GCCTGAAAGT    4320

TGTCCGTGCT GTCATTCATT TGATACCCGA TACGGCGGGG TGGGCACAGA GTATATTGAG    4380

GAAGCAGTAC AAGCGCTATT TCCTGAATAC CGTATTGCAC GGGTGGACAC CGAtGCGCTG    4440

CGCTCAGGGC ACGTGCAGCA GACGATGGAG CAGTTTCGCG CGGGGAAAAT CGATGTACTG    4500

TTGGGTACGC AAATGATAGC AAAGGGATTT AATTTCCCTA CGCTGCGTTT AGTGGGTATT    4560

GCCTGCGCAG ATACTGGACT GCACACGCCA GACTTTCGCG CCGCCGAGCG GAGTTTTGCC    4620

TTGATGATGC AAGTGGCCGG ACGTGCAGGT CGCTATGTAG ATAACGGCCT GGTCATCATC    4680

CAAACACGCA ATCCTGCGCA TCtGCGGTGG TGTGTGCGCa GCACGGGGAT TGTGAGTCCT    4740

TTTATGCGCA AGAACTTGCG CAgCGGGAGG CGCTGTGTTT TCCGCCCTTT GTGCGCCTTA    4800

TTCGGTTTGT TTTTCGCAGC AAGACGCGGC GCAAGGCTAA AGACGCCGCG TATGCGGCAC    4860

ATGCGCTTTT GACGGCGCAG ATGCCTCTGG GTGCGGATGT ACTGGGACCT GCAGCGTGTG    4920

TGGTGGCGCA GGTGGCAGGC AGCTATCGGA TGCAAATACT GCTGCGTGCC CCATCATTCC    4980

CAGTGGTGCA GCAGGTGGCG CGCAGCTTTT TAGATGAATT TCGAGCTCCG GCGGGGGTGT    5040

ACGTAGAATC TGACGTAGAT CCTGTAAATG TACTGTAGGG CGAGTAGATG TACTCCGTGT    5100

TATCCTGCTG TTTGCGTGTT TGGTTGACCG GTAGTATGCG GTGCCTGGTA TAGGTGCGGG    5160

ACGGAAAGGA GAGAGGATGT GGCACTGCCG ATTATTTTTC AGGACGCAGC gGTGGTGGCC    5220

GTTGATAAGC CGGCAGGACT TGCAGTACAG CCGGGTGCGC GGGTGCGGGT GTGCGTAgTT    5280
```

```
GACGTATTAC AGAAACAGCT TGGGGTGCGT CTGTTTCCTC TGCATCGTTT GGACAAGGAC    5340

ACCGCGGgCG TGCTGCTGTT TGCAAAAmAT GCACGGGCAG CTGCTCTGTA CCAGGGGATT    5400

TTAGGCAGCA TGCGTGTGAT TAAmGtATCG CGCACTTTGT TTTGGGCGAC CTCCCCGAGA    5460

gTGTGGTGAT ATtCGCGTTC CTATCcGTAC CGGTACGGCa GCAAGGCGGC GTCAgGTTGT    5520

gCGTGCCGCG CATACTGCAT ACCGTGTGTT GCGTGCGACT GATACGCATA CATATCTTGA    5580

ACTCACtTGC ACAGTGGTCG GACCCATCAG ATTCGTATTC ATCTGcTGCG CTAGGATGTC    5640

CTATAATTGG GGATGACAAA TACGGTGATT TCGCGCGTAA CAAGGCGTGT GCTCGTGCGT    5700

GGGGAGTAAA AAGGCTCCAG TTATTCGCAC ACAGTCTTGT GTTGCCATGT GCATGTAAAC    5760

CGCTGGTGTT GCGTGCACGT ATGCCTGTAC ACTTCCTGCG TGCTCTTGAT GCCgTTGcGC    5820

TATGATTGCC tGTAGCAGGG CATTCTGGTA rGCGGTGTGT GGTTTTGAGT TCTGCCGGTA    5880

ACAGAAAGAG TGTCGTGTGA ATTTCAATAG TTTTTCTCTA GGGTGTGTAC TGCACTCGTT    5940

GTGTTTTTGC AGGCGCGAGG GGAGGGGAGC GGTCCCCTGC TGCTGTACTG TCTGTAGGGA    6000

AGATACCGGC GCCTATTGTT ATATCGGGCT ATTTGTGCTA GAGTGTGCGA AACCGCTAGT    6060

GGGGATGGCC TATGGGTACT GTTGTTCCGG GATTCGATGA CGAGAAAGAC GAAAGTCTTA    6120

AGATGAATCT GCAAAAGATC GATGACCTTG AAGGTGGCGT CGTTGTTTTC CTCAACGGGT    6180

ACATCGATAC TTACAATTCT TCCTTTTTTC AAAAGAGGAT TGCGAAGGTT ATCGATGCAG    6240

GCTACACGCG TATTGTATTT AACTGCGCCT CTTTGAATTA TGTCTCCTCC ACTGGAATtG    6300

GTTCTTTTAC GGCGTTTCTA AAAACGGTCA AGCCtAAAGG TGGCGATATT GTTCTCCTCG    6360

ATATTCAGCC GAGGGTGTAT GAGGTTTTCC AGTTACTTGG TTTTTCTCAG TTTTTTAACA    6420

TTCGCGATTC TATTGCGGAT GCAGTTAGCC TTTTTAGGAA CAAGgTCTCa CCGyTGAAGg    6480

TGGAcACCTT TCCgAAGGTG TTTTcTTGCC CgATCTGCTC TAAGAAgTTA AAGGCGACTA    6540

AGCAGGGGCG TTTTCGTTGT TCCGAATGTA AGACGATTyT CGCCCTTGAC GCGAGCGCAC    6600

ACGTGTCTCT CGGTTAGGTG ACGCGCCTTT CTTTCTGGGG CAGGCTGGTT GGCTGCCTGT    6660

TTAGGGAGGT GTTTCGTGCT TGATGTGTGA GGTGAGGCGT TATACAATGC GGGCCGGCCT    6720

CCGGGCGGCT CGGGGAaGTC CTGTCTGTGT TGCTTCTGTA GCTCAGTTGG CAGAGCGCAA    6780

CCATGGTAAG GTTGAGGTCA GCGGTTCAAT CCCGCTCGGA AGCTTCCGTC tGTGGATGTG    6840

AGGAGGGGTG GTATGGCAAA GAGGACGGCG GTGGAGCTTA TTGCGCTTCA GTGCACTGGA    6900

TGCAAGCGGC GTAATTACAC CACTTCAAGA AACCGACGTA ACGTTCAGGA AAAGCTCGAG    6960

CTCAGGAAGT ATTGTCCTTT TGAGCGTAGA CGTGTGCTGC ATAGAGAGGC GAAGATAAAG    7020
```

```
TAGGCTGTCG TCATATCTGT TACGCACGGG GTTTTCTGGT GTTTTCCGGG GATTTGTGGG    7080

TCAGTAGCTC TAATGGCAGA GCGTCGGTCT CCAAAACCGA ATGTTGAAGG TTCGAGTCCT    7140

TCCTGGCCTG AGTGCTTTCG AAAAGGTGTT TCATGTTGAA GTTCGCAAAG TTTCGTAGGG    7200

AGTGCGTTGC CGAGTTCAGG AGGGTGGTGT GGCCTGCGCG CACTCAGGTA CATACCGCGG    7260

TTAAGGTAGT GCTCGTCTCT ACCGTTGTCA TGGCGCTTTT CCTCGGGCTT ATCGATGCTC    7320

TGTTCGTGGC GTTGCTGAGT TTCTTCTTCT GAGGGGATAG AATGGCGAAA GAGTGGTATA    7380

TTCTGCACAC ATTCTCGGGT CGCGAGGCAA GGGTGGAGCG GGCTGTCCGT ATGCTCGTGG    7440

AGCATGCGAG GATTCCAACG AACGTTATCT TTGATATAAA AATCCCTGAG GAACTGCTTA    7500

CCGAGGTGAA AGATGGTAAG AAGAGGGTGG TTAGGCGTAA GTTTTTCCCT GGTTACTTGT    7560

TGGTGGAAAT GGATTTGCCC GAGGTTGACT GGAGGATAGT GTGTAACGAG GTGCGCAGGA    7620

TTCCTGGTGT TTCCGGTTTT TTGGGTTCTT CGGGCAATGC GAACCTCAGG CGGTTTCTGC    7680

GGATGAAGCT CGGCGTATTT TGCAGAAGGC GGGGGAAATT AAGGGGGATA GGACTCCTCG    7740

TATCGCTCAG ACTTTTTTGG TTGGACAACA GGTGAGGATC GTTGAGGGGC CGTTTGCTAC    7800

TTTCTCGGGT GAGGTGGAGG AGGTGATGAG TGAACGCAAC AAGGTGCGTG TGGCAGTCAC    7860

CATCTTTGGC CGCGCTACTC CTGTGGAGTT GGAGCTAGTC CAGGTGGAGG CGCTCTGATT    7920

TTCTTCTTCC AGGGTGGAGA GTGTTGCAAT GCGCATGATT GCCTGCCGCT TACGCGTTGG    7980

TTTCGGGTGT TTTGTTGTTT TTTACGTCAT AAGGAGAGGC CAGTATGGCA GCGAAGAAGA    8040

AAGTGGTTAC TCAGATAAAG CTGCAGTGTC CTGCAGGCAA GGCGACGCCC GCGCCGCCGG    8100

TTGGGCCTGC GCTTGGGCCG CACGGGGTTA GTGCCCCGCA GTTTGTGCAG CAGTTTAATG    8160

ACCGTACTAA ATCCATGGAG CCTGGGTTGG TGGTGCCAGT GGTTGTCACC GTCTATTCTG    8220

ACAAGAGTTT TTCGTTTGTG CTGAAAACGC CGCCTGCGGC TGTTCTTATT AGGAAGGCGT    8280

GTGGGATCGA AAAAGGATCG ACGAATTCTG TTAAGCAGAA GGTTGCGCGC TTGTCGCTGG    8340

CGCAGTTAAC GGAGATTGCT CAAGTGAAAT TACCTGATAT GAGCGCTTTA ACTCTCGATG    8400

CTGCGAAgcG TAnTCATCGC GGGTACGGCA CGCAGCATGG GGGTGGAGGT AGAGCGTTCA    8460

TTATGAAGAG GGGGAAGAAG TATCGCGCTG CCGTTGCGCG TTATGATCGC GCCGAGCGGT    8520

TCAGTCTTGA CCGTGCGGTA GGTTTGCTTA AGGAAGTGAG GTATGCTTCC TTTGACGAGA    8580

CGGTGGAGGT GCACGTTAGT CTGAGGCTTA AGAAGAATCA GACGGTGAGG GATACGGTTG    8640

TGCTCCCCCA CCGTTTTCGG GCCGAGGTTC GTGTGCTCGT TTTTTGTAAA GAGGATCGTG    8700

TTTCGGAAGC GCTTGCTGCA GGTGCTGCCT ATGCAGGCGG TGCTGAATAT CTTGAGAAGG    8760
```

```
TAAAAGGAGG CTGGTTTGAC TTCGACGTGG TCGTTGCTAG TCCTGACATG ATGAAGGACG   8820

TCGGTCGTCT TGGTATGGTG TTAGGTCGCA GAGGGCTGAT GCCTAACCCG AGGACTGGCA   8880

CGGTCAGTGC GGACTTGGGG GCTGCTGTCT GTGAGTTGAA AAAGGGGCGT GTCGAGTTTC   8940

GCGCGGATAA GACAGGTGTG GTCCATCTAG CAGTAGGGAA AACGACGATG GACTCTGCGC   9000

AGATTGTAGA GAATGTTGAC GTGTTTCTGT CGGAGATGGA TCGCAAGAAG CCCGTTGACG   9060

TAAAAGCTGG TTTTGTCCGT TCGATTTCGC TCAGCTCCAG TATGGGGCCT GGGATTTGGG   9120

TTGTCCATAA GTCAGAGGAG TAGTATGGCA GTACGCGCAC GAAGGCTGCA GCCGGCAAAG   9180

GTGGCTGCTG TCGAGAGCCT TACGCGTGAT TTGGGTGAGG CTTCTTCTTA TATCTTTACG   9240

GAGTATCGAG GGCTTACGGT TGAGCAGCTG AnCCgcGTTG CGsCsCGCct GCGCGAATTC   9300

TCGTGCGTGT ATCGGGTGGT GCGTAACAAT TTTGCGAATA TCGCCTTTAC GTCCCTAAAC   9360

ATGACGGTGG GAGAGTATCT GGTGGGGCCC ACGGCCATCG CCCTAGTGGA CACGGAGCAT   9420

GCGAATGGCG TCGCGCGTGT GCTGTTCGAT TTTGCAAAGG AAGTGCCTGC CTTAGTGGTG   9480

AAGGGTGCAA TTCTTGATGG GGAGGTGTTT GACGCTTCGA AGGTAGAAGC GTATTCGAAG   9540

CTTCCTGGAA AGAAAGAGCT CGTTTCCATG TTCTTGTCCG CGCTGAATGC aACGACGGTG   9600

AAGTTCGTAC GCgTATTACA GGCTGTGATG GACAAAAGGG ATGAgGGTGT AGAAgTTTCC   9660

GTGGTGTCGG GAgGTGATTC GTCCtAgGCg GTTGTTGTAA CTTAGTTACG GGGTATGTGT   9720

TaGGCcGGTc AGGCTTCTGG GGTGCTGTCT TCCTGTCCGT TTATAGGGGT TATTTCGCAT   9780

ACAAGGAGAA GATAATATGG CGGCGTTGAG TAATGAACAG ATTATTGAgG CGATTCgGGG   9840

CAAGACCATC CTGGAGCTTT CTGAGCTTAT CAAGGCGGTG GAGGAGGAGT TTGGAGTTAC   9900

CGCGGCTGTG CCgGTAGCGC CGGTAGCGGA AGGTGGCGGG GCaGGTTCTG TAGCCGCTGA   9960

GGAGCaGACA GAGTTTACTG TTGTGCTTAA AGGACTTGCA GAACCAGGCa AAAAAATCGC  10020

GGTTATTAAA GAGGTGCGCA ACGTTATCTC AGGGCTTGGC TTAAAGAGG CGAAGGATCT  10080

GGTGGAGGGT GCGCCAAAGA CTTTGAAAGA AAATGTATCC AAGGAAGAGG CGGCAAAGAT  10140

AAAAGAGTCA ATGACCGCAG CGGGTGCGCT CATTGAGATT TCCTAGTGTC TGGTTTTTTT  10200

TGCATGCGTC CGGCGCGTCG TTGTGTGCCT CTGACACCCT TTCGTGTGGG AGGGCGTCGC  10260

GCTTTTGAGT AGAGCGTGGG CTTCTATTTC TTTTCATACT TGTTCTCGGC ATTTTGGCAT  10320

GCGGGTTGGG TCGCGTTCTC CTCACTTGAG TGGAGGGGAC GGCGTCTCCC CTGTGTGGGG  10380

AGTATTACGG TAGAGCGTGT GGTATAGGGA GCACCGTGTC GGTTCGGTGC AGCTTGAGGG  10440

GGGAGTGCAT GTCAGCACGA GTTTGCAAAA CACACAGAGT GTACGTGGGA AGGGATGTCA  10500
```

GGAATTTTAT GGACATCCCG GATCTCATCG AAATCCAGCT TCGATCTTAC GACACcTTTC 10560

TGCATGGGGC CCGGAATACA CCGTCCGGCG CCGACACCCT TATCTCCGGT ACTAGAGAGG 10620

AGCTCGGCCT CGAAGACGTG TTCAAGACTA CCTTTCCTAT CGAGAGCTCT ACGGGGGACA 10680

TGACGCTCGA GTACCAATCA TACTCCCTTG ATGAGAAAAA CATCAAGTTC TCCGAGGCGG 10740

AGTGTAAACA AAAGGGTTTG ACGTACGCCA TTCCGCTGAA GGCGCTTGTT GATTTACGTT 10800

TCAATAATAC GGGGGAGATT AGGCGCAAAG ACATTTATAT GGGAGATATC CCCAAGATGA 10860

CTGAACGCGG CACCTTTATC ATCAACGGTG CGGAgcGTGT GGTGGTATCC CAGATCCATC 10920

GTTCCCCTGG TGTTGTCTTT TCTCATGAGA AGGACAAGGA AGGACGGGAG GTATTCTCCA 10980

GCCGCATTAT TCCGTACCGG GGAAGCTGGC TTGAATTTGA AATTGATCAG AAAAAAGATC 11040

TCATCTATGC AAAGCTTGAT AAAAAGAGAC GTATCCTAGG CACCGTGTTT TTGCGTGCGT 11100

TGCACTACGA AACGCGTGAG CAGATCATCG AGGCCTTTTA CGCCATAGAA AAGACGCCTG 11160

TTTGTCAGGA TCGTGCGGAG TACGAGCTGC TCACAGgTAA GATCCTAGCA CGATCGGTGA 11220

CGGTGGAAAA TGAGCAGGgT GAAACCGGGT GTTGTACAAA GCAGGAGAGA AAATCCATCC 11280

CCATGTCATC GATGATCTGC TGCAAAACGG CATATGTGAG GTCTACATTA TTAACCTTGA 11340

AGCGGAAGGT TCGTTGCGTT CTGCGGTCGT TATCAATTGT CTTGAACGAG AGGAAATGAA 11400

GTTCTCTAAG TCGGGTGCAC AGGACGAgCT TTCGCGTGAA GAGGCACTGT GTATTGTATA 11460

CTCAGCGCTA AGACCAAGCG ATCCTATGAC CATGGACGCG GCGGAAAAAG ATTTGCAGAC 11520

AATGTTTTTC TCCCCACGTC GCTATGATTT AGGGCGGGTG GGGCGCTACA AGCTGAACAA 11580

GAAATTTCGC TCTGACTCGC CGACTACTGA GTGCACGCTC ACCCTCGATG ATATCGTAAA 11640

TACCATGAAA TTTCTCATCA GAATGTATAG CGGTGATGCA CAGGAAGATG ATATCGATCA 11700

CCTGGGCAAC CGTCGTATTC GTTCGGTGGG GGAATTAATG ACCAATACGT TAAAAACGGC 11760

CTTTTTGCGC ATGGAACGTA TTGCGAAGGA GCGTATGAGT TCTAAGGAAA CGGAAACGAT 11820

CAAGCCGCAG GATCTCATTT CCATAAAACC TATCATGGCT GCGATTAAGG AGTTCTTTGG 11880

TGCAAGTCAG CTTTCTCAGT TCATGGATCA GGTCAATCCG CTGGCGGAGT TGACACACAA 11940

GCGGCGTTTG AACGCACTTG GTCCTGGTGG ACTTTCAAGG GAGCGTGCTG GGTTTGAGGT 12000

ACGCGATGTG CACTACACGC ACTACGGTCG GATGTGTCCC ATTGAGACCC CCGAAGGACC 12060

AAATATCGGT TTAATTGTTT CTATGGCCAA TTACGCACGC GTTAACGGGT ATGGGTTCTT 12120

GGAGGTGCCG TATGTACGGG TGCGTGACGG AGTTGTTACG AAAGAGATTG AGTACCTGGA 12180

TGCTATGGAC GAGGATCGCT ACTACATTGG GCAGGATTCT ACGGCGGTAG GACCGGACGG 12240

```
GGTCATCCGT GTAGATCATG TCTCTTGTCG GCACCGGGGG GATTACAGTA CGCGTAGTCC    12300

TAaGGATATC CAGTATATGG ATGTTTCCCC CAAGCAGATA ATTTCTGTTT CTGCTTCTCT    12360

CATACCGTTT CTTGAGCATG ATGATGCTAA CCGTGCGTTA ATGGGGTCGA ACATGCAACG    12420

GCAGGGAGTG CCGCTTATTT TTCCTGAACC CCCGCGCGTG GGTACAGGCA TGGAAGAGAA    12480

GTGTGCATAT GACTCTGGAG TGCTGGTGAA GGCAAAGCAA GACGGAACGG TTGCCTACGT    12540

TTCCTCAGAG AAGATAGTGG TTTGTTCCGC CGCGGCGTCT GGGGAAGAGC AGGAGGTCGT    12600

GTATCCGTTA CTTAAGTATC AGCGGACAAA TCAGGATACC TGTTACCACC AGCGGCCAAT    12660

AGTGCACGTG GGAGATCGGG TACAGGTAGG AGATGCGCTT GCAGACGGTC CTGCAACGTA    12720

TCGAGGGGAG CTTGCGCTTG GCAGAAACAT TCTAGTTGGT TTTGTGCCGT GGAACGnTTA    12780

CAACTACGAG GATGCCATTT TGATTCTCA CCGGGTGGTA AAGGAGGATA TGTTCACCTC      12840

GGTTCACATC AAAGAATTTT CTACTGAGGT GCGTGAAACC AAGCTGGGTT CTGAACGAAT    12900

GACGAATGAT ATCCCGAATA AGTCTGAGAA GAATCTGGAT AATTTGGATG CAGAGGGGAT    12960

CATTCGTATT GGGTCAAAGG TGCGTGCGGG AGACGTGCTT ATCGGAAAGA TTACGCCAAA    13020

AAGCGAGTCT GAGACGACGC CAGAGTTTAG GCTGCTGAAT TCTATTTTTG GGGAGAAGGC    13080

GAAGGAAGTG CGTGATTCTT CTCTACGTGT GCCGCATGGA GTTGAGGGTA CAGTCATTGA    13140

CGTGCAGCGA CTCAGGCGTT CGGAGGGAGA TGATTTAAAC CCCGGGGTGT CAGAGGTGGT    13200

GAAGGTTCTT ATCGCTACCA AGCGTAACTG CGTGAAGGGG ATAAAATGGC CGGTCGCCAC    13260

GGTAACAAGG GTATCGTTGC GCGCATCCTT CCTGAAGAAG ACATGCCGTA TCTGGATGAT    13320

GGTACCCCGC TTGATGTCTG TTTGAACCCG CTCGGTGTAC CTTCTCGTAT GAACATAGGA    13380

CAGATTCTTG AATCTGAATT GGGACTTGCG GGGTTGCGGC TTGACGAATG GTATGAGTCT    13440

CCTGTCTTTC AATCTCCAAG CAACGAGCAG ATTGGGGAAA AGTTGATGCA GGCAGGTTTT    13500

CCGACTAATT CAAAAGTGAT GCTGCGTGAC GGACGCACGG GGGATTATTT TCAAAACCCT    13560

GTATTTGTGG GGGTTATTTA CTTTATGAAG CTTGCGCATC TAGTGGATGA CAAAATGCAC    13620

GCCCGCTCTA CAGGTCCATA TTCGCTTGTG ACGCAGCAAC CCTTAGGGGG TAAAGCGCAG    13680

TTTGGAGGGC agCGTCTCGG GGAAATGGAG GTGTGGGCGC TTGAArCcTA CGGCGCGGCG    13740

AATACCCTGC AGGAGTTGCT AACGATTAAA TCGGATGATA TGCACGGGCG TTCTAAAATT    13800

TATGAGGCAA TTGTAAAAGG GGAGGCTTCG TCTCCTACCG GTATTCCTGA ATCTTTTAAC    13860

GTGTTGGTGC AGGAGCTGCG GGGACTTGCG CTCGACTTTA CGATTTACGA TGCGAAGGGC    13920

AAGCAGATTC CGCTCACTGA GCGCGATGAA GAAATGACGA ATAAGATTGG CTCTAAATTT    13980
```

```
TAAGGGGTGC aGGGAATGAA GGATATCCGG GATTTTGACA GTTTACAGAT AAAGCTTGCC    14040

TCCCCTGATA CCATTCGGGC ATGGTCCTAT GGAGAGGTGA AAAAGCCTGA GACAATTAAT    14100

TACCGCACGT TGCGTCCTGA ACGTGAAGGG CTTTTTTGTG AACGCATTTT TGGTACTACA    14160

AAGGAATGGG AATGCTTTTG TGGAAAGTTT AAGTCAATTC GGTACCGGGG TGTTATCTGC    14220

GATCGGTGCG GGGTGGAGGT AACGCATTTC AAGGTTCGCA GGGAGCGCAT GGGGCATATT    14280

GAGCTTGCAA CGCCTGTTTC TCATATTTGG TACTACCGTT GTGTACCAAG TAGAATGGGT    14340

TTGTTACTCG ATCTACAGGT GAnTCgCAtG CGTTCTGTTT TGTACTATGA GAAGTACATA    14400

GTTATAGAGC CgGGCGACAC CGATTTAAAA AAGAATCAGT TGCTCACTGA AACTGAGTAC    14460

AATGaCGCGC AGGAGCGCTA CGGTGGCGGC TTTACGGCGG GAATGGGAGC GGAGGCTATC    14520

CGTACCCTTT TGCAAAACCT TGACCTTGAC GCGCTTGTTG CACAGTTGCG TGAGAAGATG    14580

ATGGAGAAGG GTGCGAAAAG CGACAAACGC TTGCTGCGTC GCATAGAGAT CGTAGAAAAC    14640

TTTCGGGTGT CGGGAAATAA GCCGGAATGG ATGATTTTGA GCGTTATCCC GGTGATCCCG    14700

CCTGATTTGC GTCCTATGGT GCAGCTCGAC GGAGGGCGTT TTGCTACCTC AGATCTCAAT    14760

GACCTGTATC GGCGTGTGAT CCACCGCAAT AGCCGTTTGA TTCGGCTCAT GGAACTGAAG    14820

GCGCCGGATA TCATCATTCG GAACGAAAAG CGCATGTTGC AAGAGGCAGT GGACGCGCTT    14880

TTTGATAATT CTAAGCGCAA gCCCGCGATT AAAGGTGCGT CAAACCGGCC GCTTAAGTCT    14940

ATTTCTGACA TGCTCAAGGG GAAGCAAGGG CGTTTTCGCC AGAATCTTTT GGGCAAGCGG    15000

GTCGACTATT CCGGGCGTTC GGTTATCGTA GTGGGGCCTG AACTTAAGTT GTGGCAGTGC    15060

GGGTTGCCTA CAAAAATGGC GCTTGAGCTG TTTAAGCCCT TTATTATGAA AAAGCTGGTT    15120

GAGAAAGAAA TTGTCTCGAA CATCAAAAAG GCAAAGATGC TCGTGGAACA AGAGTCGCCG    15180

AAGtATTTTC GGTGTTGGAT GAAGTGGTAA AAGAGCATCC AGTTATGCTT AATCGGGCGC    15240

CGACATTGCA TCGATTGGGC ATTCAGGCTT TTGAGCCGGT GTTGGTGGAG GGGAAGGCGA    15300

TTCGTCTTCA TCCGCTTGTG TGTAAACCTT TTAATGCTGA TTTTGATGGG GATCAAATGG    15360

CGGTGCATGT GCCGCTGACG CAGGCGGCAC AGATGGAGTG TTGGACGCTC ATGTTGTCGA    15420

ATCGCAATTT GCTTGACCCT GCAAATGGGC GCACGATTGT GTATCCATCT CAGGACATGG    15480

TTCTGGGTTT GTATTATCTG ACAAAGGAAC GCTCTCTGCC GGAGGTGCTC GTCCTCGCCG    15540

TTTTTCCTCG GTGGAGGAGG TAATGATGGC TGCGGAAAAG GGGGTAATCG GCTGGCAGGA    15600

TCAGATTCAA GTGCGATATC ACAAATGTGA TGGTCAGCTT GTGGTCACTA CCGCAGGAAG    15660

ACTTGTGTTG AATGAGGAAG TTCCCGCAGA GATTCCTTTT GTCAACGAAA CGCTTGATGA    15720
```

```
CAAACGCATC AGGAAATTAA TTGAGCGGGT GTTCAAGCGT CAGGATTCTT GGCTTGCGGT    15780

GCAGATGCTC GATGCACTGA AAACTATCGG TTATACCTAC GCGACCTTCT TTGGTGCAAC    15840

GCTCAGTATG GACGACATCA TCGTGCCTGA GCAGAAGGTG CAGATGCTCG AAAAGGCGAA    15900

CAAGGAAGTG CTAGCGATTG CGAGTCAATA CCGCGGGGGG CACATCACGC AAGAGGAGCG    15960

TTATAATCGC GTCGTTGAGG TGTGGTCTAA AACAAGTGAG GAGCTCACTT CGCTCATGAT    16020

GGAAACACTT GAGCGCGACA AGGATGGATT TAATACCATT TACATGATGG CTACCTCAGG    16080

TGCGCGCGGG AGTCGCAATC AAATCgCCAA CTGGCGGGAA TGCGTGGCTT AATGGCAAAG    16140

CCGAGTGGGG ATATCATCGA ATTGCCTATT CGTTCTAATT TTAAAGAGGG ACTCAATGTC    16200

ATTGAGTTTT TTATTTCTAC CAACGGTGCA CGCAAAGGGC TCGCAGACAc TGCGCTAAAG    16260

AcCGCTGATG CGGGGTATTT GACACGTCGT CTGGTTGATA TCGCGCAAGA TGTGGTGGTG    16320

AACGAGGAGG ACTGTGGTAC CATCAATGGC ATTGAATATC GCGCGGTGAA GTCCGGCGAT    16380

GAGATTATTG AATCGCTTGC TGAGCGCATC GTAGGAAAGT ATACACTTGA ACGTGTAGAA    16440

CACCCCATCA CCCATGAACT GCTGCTCGAT GTGAACGAAT ACATCGACGA TGAGCGTGCA    16500

GAAAAGGTGG AAGAAGCGGG CGTGGAGTCA GTGAAGTTGC GCACCGTGCT CACGTGCGAA    16560

TCTAAGCGAG GAGTGTGTGT GTGCTGCTAC GGGCGGAATC TTGCACGCAA CAAAATTGTA    16620

GAAATTGGGG AGGCGGTTGG GATTGTAGCC GCTCAGTCCA TTGGTCAGCC GGGTACGCAG    16680

CTGACAATGC GCACGTTCCA TGTTGGGGGT ACGGCAAGCA GTACTACGGA AGAGAACCGC    16740

ATCACGTTTA AGTATCCCAT ACTGGTAAAG AGTATTGAGG GGGTGCATGT GAAAATGGAG    16800

GATGGCTCTC AGCTGTTCAC GCGTCGGGGG ACGCTCTTTT TTCACAAAAC TCTGGCAGAG    16860

TATCAGCTTC AAGAGGGTGA CAGCGTGCAG GTGCGTGACC GCGCGCGGGT GCTAAAGGAT    16920

GAGGTTCTCT ACCACACCAC CGATGGGCAG ACGGTGTACG CTTCGGTGAG TGGTTTTGCG    16980

CGTATAATCG ATCGAACCGT GTACCTGGTA GGGCCTGAGC AAAAGACGGA AATTCGCAAT    17040

GGTTCTAATG TAGTAATCAA GGCAGACGAG TATGTGCCGC CCGGAAAGAC CGTGGCTACG    17100

TTTGATCCGT TCACTGAACC TATTTTGGCA GAGCAGGATG GCTTTGTGCG GTACGAAGAT    17160

ATTATTTTGG GCTCTACGCT CATCGAAGAG GTAAATACTG AAACGGGGAT GGTGGAGCGC    17220

AGGATTACGA CGTTGAAAAC AGGAATACAG CTTCAACCGC GGGTATTCAT CTCTGATGAG    17280

TCGGGGAATG CGCTGGGTTC GTACTACTTG CCAGAGGAAG CGCGCTTGAT GGTTGAAGAA    17340

GGCGCGCaGG TGAAGGCGGG TACGGTCATT GTAAAACTGG CAAAAGCAAT TCAAAAGACA    17400

TCGGATATTA CGGGGGGGCT GCCGCGTGTT CTGAATTAT TTGAAGCGCG GCGCCCTAAG    17460
```

430

```
AATGCGGCTG TCTTGGCACA GATTTCTGGG GTTGTGTCGT TCAAAGGACT GTTTAAGGGT    17520

AAGCGTATTG TCGTGGTGCG TGACCATTAC GGGAAGGAAT ATAAGCACCT CGTGTCCATG    17580

TCGCGTCAgC TTTTAGTACG TGATGGAGAT ACGGTTGAGG CAGGCGAACG CTTGTGTGAT    17640

GGTTGCTTTG ATCCCCATGA TATCCTGGCA ATTCTGGGTG AAAATGCTTT GCAAAACTAT    17700

TTGATGAATG AGATCCGTGA CGTGTATCGT GTGCaGGGTG TTTCAATCAA TGACCAGCAC    17760

ATTGGTTTAG TGGTGCGGCA AATGCTACGA AAGACAGAGG TTGTCTCGGT TGGGGACACG    17820

CGTTTTATCT ACGGGCAACA GGTGGATAAG TACCGTTTTC ACGAAGAGAA CCGTCGGGTT    17880

GAAGCGGAAG GGGGGCAGCt GCGGTTGCGC GCCCAATGTT CCAGGGTATA ACGAAGGCGG    17940

CGTTGAACAT AGACTCTTTC ATATCTGCGG CATCTTTCCA AGAAACGAAC AAGGTGCTCA    18000

CCAATGCGGC GATTGcAGGC TCTGTTGATG ACTTGTGTGG GTTGAAGGAG AACGTCATTA    18060

TAGGGCACTT AATTCCCGCA GGTAcGGGGA TGCGGCGTTA TCGTCAGGTG AAGCTGTTTG    18120

ACAAGAACAA GCGGGATCTT GATGTGCaGA TGGAGGAAGT TATCAGGCGT AGAAAACTTG    18180

AAGAGGAGGC GCTTGCCCAG GCAGTTGCGG GTATGGAAGG GGAACCTGAA GGCGAAGCGT    18240

GATGGATTGA CCTGGTTTGG CTATTCTGAG TATCCTAGTC CGCGTGTGCT GTGTGCGGCA    18300

AGGTTTACGG TGTTGAGGAT TTTTTTGGGG AAGTGAGCGA AAAGAATGCC GACAATTAAT    18360

CAATTGACGA GGATAGGGCG TAAGGCGGTT TTTTCTCGTA CGAAGAGCCC TGCGTTGcAG    18420

GCTTGTCCgC AGAAGCGCGG AGTGTGTACG CGTGTGATGA CAGTTACGCC AAAAAAGCCG    18480

AATTCTGCTC TGCGTAAGGT GGCGCGTGTG CGTCTAAGTA GCGGGGTTGA AGTGACGGCG    18540

TACATTCCCG GGATTGGGCA TAATTTGCAG GAGCACTCGA TTGTGCTGAT TCGCGGTGGA    18600

CGTGTGAAAG ATTTACCTGG AGTACGTTAT CATATTATCC GGGGGGCCAA GGACACTCTT    18660

GGCGTGGTGG ATCGTAAGCG CGGTCGTTCA AAGTACGGGG CTAAGCGCCC TCGCGCGTAG    18720

GGGCTGGGGA GAGGAGTTGG TATGGGGCGG AAGCGACGGG TGTCGCGTCG GGTACCGCCG    18780

CCTGACGCGC GGTATAACAG TGTGGTGTTG GCGAAtTTAT TTGTCGAATG ATGCTGGCGG    18840

GTAAGAAGGC AACTGCGGTG GGTATTATGT ACGATTGTCT TGAACGTATT CAGCAAAGGA    18900

CTGGTGAGGA GCCTCTTCCG GTGTTCACAA AAGCGTTAGA GAACGTAAAG CCTGCAGTGG    18960

AGGTTAAATC GCGGCGGGTT GGTGGTTCTA CCTATCAGGT GCCGATGGAA ATTCGGGAAA    19020

CGAGGCGTGA GGCTTTAGGT ATGCGCTGGA TTATCGGTGC AGCACGCAGG CGCACGGGAC    19080

GTGGCATGTC GGAGCGACTT GCAGCAGAGA TCCTTGATGC GTACCACAGC ACGGGAACTG    19140

CCTTTAAACG TAAAGAGGAT ACGCACCGCA TGGCAGAGGC CAATAAGGCT TTTTCGCACT    19200
```

ATCGCTGGTA GATACGCGTC TCTTCCTGGG GCGTTTGTTG CAGGGGCGGT GTCTGCCCTT   19260

GGCAGGGGTG TTTTTGCCCT CGTCCTTTCT CTTGATTCAT CTGGACGTCG GTTTTGGGTG   19320

GCGTGCTCTT GTGCGCCTTA TCAGCATAAA CGGAGGGTCC ATACGGTGGG GGGGCTACTC   19380

TCGGATCCAC ATAATTTTGC GCGCGCGTGT GCCCTCTTTC GTGGAATTTT CCGCAAGGGA   19440

AGAGCGCTCG GGGGTGGTTC GCGCAAAGCT TCAAGTGCCC TGT   19483

(2) INFORMATION FOR SEQ ID NO: 43:

     (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4724 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

CCTTTTTTCG ATCTGTCCAA TATGAGTGGT TGGACGAGCG GACATTTTGT GGAAATGGAA   60

TCCGCTCTGT CTGAGTATAA AAAGTCAAAA AAnCCGCTCT ACGTTTTTTC TACCTCTTAC   120

AGTTTGGCTG ACTATTACAT CGCCTCTTTT GCTGATGAAA TTATCCTTGA TCCGATGGGG   180

TCTGTGGATC TGTCGGGCTT TTACACGGAA ACTCTCTTTT ACGGAGGTAT GGAGGAAAAG   240

ATTGGGGTGC GTTGGAACGT CGTGCATGCT GGGGTGTAnA AGGGCATGGC TGAGATCTTT   300

TCTAGGAAGG ATTTTTCTCC TGAGGTTCGC AGAAATTATC AGTCTGTATT TGCGCGTCTG   360

TGGCAGCAGT ATCTCAGTGA TGTTTCGCGT AATCGAGCAC TAGAGGTGCA GCATCTTGCC   420

CGTTACGCGG ATCGTCGCCT TGAGCTCCTG CAGAAGTATA ACGGAGACGG TGCGCGCACC   480

GCATTGGCGG AAAAGTTAGT AACGCGCGTA TGTTCCTACG ATGAAGCTGG CGTTGCGCTC   540

AAATTTTTAA AAGAAGACGA CTACGAATCT GCAAAAAATT TCGTTGGTCT AGACGATTAT   600

AATCGTGACC GTGCACAGCG GCAGGTGCAG GATCAGGTGG GGATTATTCA TCTTGCAGGA   660

CCGATTGCTG CACACAGGGA TACGGAACTC GGCGGAACGA TCAGCGACGA GGTTAGTGCT   720

TTGTTGGATG TCGCGATGAG TGATCCGGAT ATTAAGGCAG TAGTGTTGCG TATTGATTCC   780

GGTGGGGGAG AGGTGTTTGC TTCTGAACGT ATCCGCCGCG CGCTTGCGnG GCAAAGCGT   840

CGAGGCAAGA AGCCAGTGAT AGTATCGATG GGTGCGATTG CTGCGTCTGG TGCGTACTGG   900

GTTGCTTCTG CAGCCGATTA CATCTTCGCA TCCCCCTATA CCATCACTGG TTCCATAGGG   960

GTGCTTTCGG TACTACCGAC ATTCGAAACG TTTTTAGAGC GATATGCGGG GATCACTGTC   1020

GATAGCGTAC AGGTGCACGG CGTTCGCCAA CCTTCTTTGC TCAGGAGTGG AACGGCTGAA   1080

```
GACACCGCGC GCATGCAGCT TGATGTGATG GCGACGTATC GTACTTTTCT TTCGGTTGTT    1140

TCTGCCGGGC GTAACCTTAC CCTTGATCGG GTGGCGGCGG TTGCAGAGGG TAGGATTTAC    1200

GCGGGGGAGG ACGCAtTTCC CTAGGCTTGG TTGATGCGCT AGGCGGACTA GATGAAGCGG    1260

TAGCACATGC AGCGAAAGAA TCACATTGCA GGCAGTATTC GGTGAGAGTT TTGAAGCGGA    1320

CCsCACGTAC GGTGAAGAAT TTCTGCAGTC CCTGTGGGAT GTCCTGCAGA AACGAAtCTT    1380

GCTTTTGGAG AGCGTGTGAT CATTGGAGAG TTACTCCAGC TTGACCTAAG CAAGGGCACC    1440

TACGTATATG AGCCGCTGCG CTTGCATTGG CGTTGACGGG CACTGCTACG CTTGATCGAG    1500

CGCACGTnGT TTGCTACGGT TGGCGCCGGT TTTTGGGGAT GTAGCTCAGT TGGTTAGAGC    1560

GCTTGCATGG CATGCaAGAG GTCAGGGGTT CGATTCCCCT CATCTCCATC GCCGTGTGTG    1620

AGGGAGGGGG TGTGTCTGAT TTAGGTTTAG ATCCGGATCT GTTAGCTCTG CTGCAAGATA    1680

CGCCGCAGGt GTGCCGTCTG AGCATTCTTC TGCAGGGAAG GGTACAGCGA TGTCGCCTAc    1740

CGGGACGCGA GATCCGAGTG ACGTTGATCT TTCTGAGCGT AgTTTTCCCT TGGTTACTGA    1800

GTTTCAAAGC AAGACCCCGC ACCAGTTTTT TGAGTCAGCA GAGTTTTATA AACGTGTCGT    1860

TTCGGATGAG TTGGAAGTTG GGCAGCGTGC GCATGCGGCT TTGGCGCGCT ATTTGTCCAC    1920

CACTGACTTA AAGGATCGCT CTGTGTGCCG GCAGCAGCTT ATTAGCAGTT ACTGGCAATT    1980

AATGGCACAG ATATCGGGGA AAATCGGCGG TGGGTCGGCG TGCATGGAAA AGCGTTACGC    2040

ATTGCGCTAT GGACTGTTGC TTCCTACCTT GTTGACCGCA TCCCAGAAAG ATATCTTCGC    2100

GCGGATTATT GAGACGAATA GTTTGCAGCA GCCTCTTTAT TATCTGGATG AATGGCTGAT    2160

TGCGATTGGT TCTGGAAAGG TTCGCCCTTC AAGCACCGAC GAAgTGCAAG TAAAAAGGAA    2220

AGACGATGTC GCACGCGTAC GGCAGGCGTA TGATAAAGCG TGCGGGCAGT TGCAGAGTTC    2280

TGAGCGTCTG TTGCAGGTGA GGTCGGCGGA gcGTGCCCGT GTGGAAGAGG AGGTGAAGAA    2340

CAGAATTTCG CGTCTTTTCG TGCACGAATC CATTGAAGGT CTCCCTGGGG TGACAGCAGG    2400

TTTCAACGAG GCGCAGAAGC AAGGAATCTC GGAGATCCAT GAATTGTTAA AAAAGTTGTT    2460

GGGTATAGAT CGGGAGTTTA ATGGGTTATA TGCGGGCTAC CGCGCTTCAC AAGACGCAgT    2520

GCATTCCCTG CGAGAGAAAC TAGATGCGCC CAATGCGGAG AACAGTTCAG CAGTGAGTAC    2580

GGAGTACGAT aCCGTGCGCC AAATGATAAA GATGAGCTGC GGGCGCCAGG GCAACCATTT    2640

CCCCCTCTTG TCCAGAGAGT ATTTCCGTTC TGCGGAGCAT GAGATTGGCA CGCGGGAAAA    2700

TGTATTGAAA ATTATGGCTT GGATTGAAGG TCTGGATCCG GAAGCGTATT GCCGTCAGTA    2760

TAAGCAGCAG GTAAACAGGA TTCCGCCATT CGTGGTGCTG TTGCCTTCTT ATGGGGACAT    2820
```

```
AGGATTTTGT TGGGAGCCGT TTGATCGTTA CAATCGCGTG ACAAGCCGTG GACGCGTTGC    2880

GtGCCTATGT ATGGAAGGAG CTTGAAGCTT GCAGTTATTA CCGCGACGGC GGATTTACGT    2940

TGGCAGGTTG CAAAGGAAAA GGCTTCGTAT TACTGGATGG AAGAGGGCTT GACGGGGAAT    3000

TATTATCAGT GGTTTCAACC CCAAAAATTA AGGGGTGATG TAAAGGAGTA TTTTATTGCC    3060

GATTACACGA CCTGGCTCCT GAAGGAAAGC GAGGGCATCC AGAAACTGGA CAAAGAGGTC    3120

CGCAATGTCT TTTGGCGCTA CATCCCCTTT CCCCAAAAAA TCAAAGACGA ACTCAAGACA    3180

AAGTCCTTTG TGTACCAAGA GCTTTGTCAG AAGGACGCCA ATCGCCAGGT ATCTGACGGC    3240

TATTGATAGT TTCTCCTGAA TCGGTTGGTG TCCTGTCATG AGGGGATAGC TTGTGCGCCG    3300

GTGTCGGGTG TTCGTTGACC GAGAAGGGTC AGGGTGTTTT TnAAGCTtys CTCTCGCGCG    3360

ATTGATGGGC AAGTCTACTG CAAGCAGGCG TGCGAGGTAG ATCCCATAGT GAGGATGATC    3420

CTCAATCAGT GAGATGAACT TCATCTTTGA TATCTTTACC AGTGTGCCTT CGCCTACTGA    3480

TACAATCGTT GCAGAGCGCC GGTTGTTGAG CAAGAACGAC ATTTCCCCGA TGAATATGTC    3540

TGATGGGGTC AGCATGGACA TGAACTTGTT ATCCACGTAC ACTGCGAATT TCCCCGACGA    3600

AATGTAGAAA AGGGAACTGG ATTCTTCGTT CTGGTAACAT ACCACCTGGG CATCCCGGAA    3660

CGTTAGTACC TGTTGGTTTT TCAAAATTGA AGGCACAAGG TTCGCGACGT TTTCCTGGTT    3720

GTCAATTTCA AACGTCACCT CGTTGCCCAC GTCGTTATAG CTGAGCCTTT TGACAAAAAT    3780

TTCTGTCATT TTTATGCCCA TACCGTGTAG ACCAGGTTTG CACGCGCTTG CCATGCGGCT    3840

TTTCCAATCA AAGCCTGTGC CTTCGTCACG AATGGTAATG CGTGTACGCT GCAGTGTAAT    3900

GTCATAGGAA ATATAGATTT TTTTCgCGCT AATGCGCGGG TCCtGcTTGC GCAGAGCAAT    3960

CAAATCAAAG ATATCCTTGC GCTGTTCGAG CCaCTCTGTT TTTTCGTCGT AGCTAATGCC    4020

GCAGTTTCCG TGCTCCAGTG CATTGAGTAA CAGTTCCATC ATTGCGCCTT CAAACGAAGT    4080

GCGTTCAAGC TCATTGATAC GGTTGGTATT GTACAGGTAC GAACTAATCA AGCTGGCGTA    4140

AAAGGTAATC TCAAAGGAAT CGGTGTCGCA GATAAAATTT CCCTGCTCGT GTCCATGTGC    4200

TTGGTGCACG AGGCTGCGAC TAGAAAGAAA GTGTCGGTTT CTGTCCACGA TTCGCACAAC    4260

CTGGGAGGCA TGCGCTTCAA ATTCCTGCCG CGTGGAAACT GAAAGGAAAT TCGGGTCTTT    4320

GCGATTTACG ATTTTTATTT TTTCTTCCAT CGAGTTAGTG ATAGCAATCA CCCCACCAAA    4380

TAGAAGCCAA GGATCATCCT TTATAATTTT TAAACACGCT TCGCTGTCGA CGTTTGGGTC    4440

ACCAAAATCA ATAATCTTAA TCTCAGGCAT CTCGAAGCGA AAAACGGATG CTATCTCATT    4500

CAGACGAGAG AGCGTCTGAA TGTGTATATC CACGCGTTCT CCAGTACACG CACCGTTAAg    4560
```

GCAGATATGG TAGACGTAAC CGTACTGATA AGAGGTATTT gCTCATACTC ATAATCCTTg   4620

TTATAGAAAT CGAGCCACGG TAATCATCGG TTGACTTATC ATcGAGAATG AGATCTGGcT   4680

ACGCATTAgG TATATCGTGT GGGGGCATGC GCnTGGGAAC AGGC   4724

(2) INFORMATION FOR SEQ ID NO: 44:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 14822 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

TAGCCTGCCG TGGCGCACCC CTGCTTTGCT CCACGCCGCG CTCTTTACGT TCCCCGCACA   60

CATGCGCTAC ACTCCCCGCC ACCGCCGCAG GcAGGGCCCC GTGTTACAGG ACCTATCCGC   120

AAATGCCCGT AAGTACTGCT CGAGCTCGGT CAGACCGTTG CGCGTGAGGC TCGCGCCTCG   180

CTCCTTCAGC CAGAGCACAT TCTGCTCGCC CTCATTCAGC ACAAAGTAGG CCGCGGCTAC   240

AAGCTCATCG AAAAACTCAT TGAAGATGTC GCTACCGTCC GCCTCATCCT CGAGCAACAC   300

GTCCTTACCA ATGAGGGAGA CGTCGCCAGT CCCCAGGACC TGCCCGTCTC AGGACGCGTC   360

AAACACTTGC TCGACATCGC AGCAATGGAA GCACGCTCCc TGCGGTGCGC TTACATCGGT   420

ACCGAACACC TCGTTATCGC CTTTGCCCGA GAGGAGCAAA ATCCTCTCTT CCAAAGCCTC   480

ATCCGAGAAG GACTCTCGCT CGATGACCTG CGAAACGCGA GCATTATATC CTCACCTCAT   540

TCTGATACCA CCCGCACCCG GCTCGAGCGG AAAGTTGCAA GTGTCCTTGA CGAATACGGC   600

ACCGACCTTA CCGAACGCGC GCGCGCCGGC GCCCTCAATC CGGTCATCGG ACGAAACAAA   660

GAAATTACCC GCGTCATTCA AATCCTGTGC CGGAGAGGAA AAAATAACCC GGTGCTCATC   720

GGAGAGCCAG GTGTCGGGAA AACTTCCATC GTTGAGGGGC TCGCGTACGC CATCGTTCGG   780

GAGGAGGTCC CGCACATCCT GCTGCACACC CGCGTCGTTT CCCTAGACCT TGCCGCCGTC   840

ATAGCAGGAA CAAAGTACCG CGGCCAGTTT GAGGAGCGGC TCAAACGCAT TATTAAGGAG   900

GTGGAAGAAA CTGAAAAAGT CATCCTTTTC ATCGATGAGC TGCACACACT CATCGGAGCA   960

GGAGGCACGC AGGGGTCTTT GGACGCCGCC AACATGCTCA AGCCGGCCCT TGCACGCGGA   1020

CAAATCCAGT GCATTGGGGC AACAACCCTG GCAGAGTATC GCCGTTACTT TGAAAAAGAC   1080

GCAGCTCTCA CCCGCCGATT CCGATCGGTG CTCGTGCGTG AACCGAGCTT TGAAGAAACC   1140

TGCACTATTT TACGCAAAAT AAAATCACAC TACGAACGAC ATCACCAGGT GATATACCAA   1200

```
AGCGATGCGC TTGAAAAAAT TGTTGAGCTT TCACGGCGCT ACATCCCTGA GCGGTTCTTT    1260

CCAGATAAGG CAATTGaTCT TATGGATGAA GTAGGAGCCA TGAAACGGGT ACAACAGCGC    1320

GCGGATACGC AGGTATTGCG TTCCTTTTCC ATAAAAGTTG CTAATCTTAC CACAGAGACT    1380

GAGCGCGCCA TTGCGCTTGA AGATTGGGCG CGCGCGCGTT CCTTACACAC CGATGTGGTG    1440

CAGCTGcGCA GACGGCTCCA CGCGCTGAAG GTAGAGTGGA GCGCGCGCGA AgyGcgTCTA    1500

TCTTTGcAGA AGATGTTGCA CAGGCTGTCT CTCTCATGAC CGATATCCCG GTACATTCGC    1560

TCGAAGGGGA TGAGCTGTGC CGCTTTACCA ATATCGAACG GGATCTTTGT GCCACCGTGC    1620

GTGGGCAGCG CGAGGCCATT GCAACGCTCG CGCGCGCTAT CGTACGCGCG CGTGTCGGCA    1680

TCTCTTCAGA CACGCGCCCC ATTGGCTCCT TCCTGTTTCT TGGACCGACC GGTGTAGGCA    1740

AAACGCTCTT GGCAAAGACA CTCGCGGAAT TTCTTTTCGG TTCAGCAGAC GCGCTCATCC    1800

GCATTGACAT GAGCGACTAC ATGGAACGCT ACAACACCTC ACGCCTCATG GGAGCACCGC    1860

CTGGATACGT GGGATTTGAA AATGGCGGTC TACTTACCGA GCGCGTACGG CACCGCCCTT    1920

TTTCTGTCAT CCTTCTGGAT GAAATTGAAA AGGCGCATCC AGATGTCTTC AATGTTCTCC    1980

TCCAGGTGTT AGAAGAAGGA GAGCTGCAAG ACAACCTGGG GCACACGGTG AACTTCCGCA    2040

ACACTATCAT CATCATGACC AGCAATGCAG GCACACGCGG CCTGGGGGAA AACGTTCCTG    2100

GCTTTCAAAC CGCACGCGCG CGAAACATCG AGTACCGTCA GcTGCGCGTA CAGGCCcTCC    2160

GGGAAATAAA ACGCATCTTC TCTCCGGAGT TTCTCAATCG CGTTGACGAG TGCGTAGTGT    2220

TTGCTCCGCT TGAGCGAGAG ACCCTGCAGG AAATTTTAGA ATGCGAACTG AAGAAGCTCG    2280

CAGAACGCCT ACGCGGTAAA GATATTGTGC TGCGCTACAG CGCGGCTGCA AAGGCCTACT    2340

GTCTTGAACA CGGCTTTGAC CCATTCTTGG GCGCACGCCC CtGCGCCGCG TATTGCAGCA    2400

AGAAATTGAA AATGAGCTTG CGcTGCGCAT GATTCACGGA ACGTTGCGCG CAGGATCGTG    2460

CGTGCACATA GACTCAGACG GCGCGCGCCT CCACCTTTCT ACCGAAAAAA GTTACCTGAC    2520

GCTGCATCCC CAAGAAATAT AACTAATCAG TCACACGCGC CCGTATCTCC CGTACCTGCA    2580

GGTCACTTTC CCACACAGAG CTTCTCAAAC AGCGCATCTA GGATATCTTC GCTGTGCACT    2640

TCTCCAGTAA GCGCCCCACA ATGATAGAGC GCCTCTTCCA GATCGTGCAC CACTGCATCC    2700

AACCCGAACC CACGTGCATA CGCCTCCTGT GCATGCTCCA ACGCCTGCAC TGCGGCGTCT    2760

ACCAATACGT ACTGGCGTTC TGAGCCAAGA GAAAGCTCCT CGTACGGCAC CTGACCGCCG    2820

TGCAGCAGGT GGAGTGTCTG TGCACGGAGC GCGTCCAACC CCGCGTGAGT CTTTGCGCTT    2880

ACACACACGA ATGCgCGCGG CGCACGATCC CTCACCTCCC CGTTCTTTCC CCCTGCTAAA    2940
```

CACTGCTCCC CCGCCCCGCG CGCGTCCTGA CTGCGCGCAC ACGACAACAC CGGTGCCGAT   3000

ATAAACGGCT GCACTGCCTG ACACACCTGT ATGCGCTCAG ACATAGACAT CAAATCGTTG   3060

TGGGTAACTA CCACTACCAA GGGTACTGCA CAGTCCGAAA GAAAAGCGCA ATCTGCAGCC   3120

TGCACACCTG CACGTCCATT AATAATGTAA AAAACGCAAT CTGCTCCCTG CAAGAGTTGC   3180

TCGCTGCGTA CCACTCCCTG TGCCTCAATA GGATTGTCAG TTACTCGTAA GCCTGCCGTA   3240

TCACACAGAC GCACTGGAAT GCCCGmTAGA TCAAGGTCTG CTTCAAGCCA ATCGCGCGTT   3300

GTACCCGGAA CGGACGAAAC GATGGCACGA TCCTGTCCTA AAAGAGCGTT GAAAAGAGAT   3360

GATTTACCCG CATTTGGACA ACCGCCGAGC ACGATGCGCA CTCCCGTTCG CTGCAGCGCA   3420

CGCTCCTGCC AGCAGGCACG GAGCCTGCGC AGACGTTCTA CCAACGGTTC AAGTTCACGC   3480

ATATCGATAT CGTGCACACG CGTTTCTTCA TCTTCCGGAT ACTCAATTTC CCCCTGAAGC   3540

GTGGCTGAAA ACGCGAGTAA CGCACGGGTA AGCGCTGCTA TCTCCTGCTG CAGCGCACCT   3600

GAAAGtGmAA CACCGCTTGC TGcTGCGCCG CACACGTGCG TGCATCAACT AGTGACTGAA   3660

TCGCCTCAAT ACGCGTCAAA TCCCTTTTAC CATGAAAGAA TGAACGAAAA CTAAATTCAC   3720

CTCGCTGGGC GGCACGGAAC CCnTGCGCAA GACAGAGCCG ATACACAGCC TGTACGGTAC   3780

GCACGCCCCC ATGACAAATA ATTTCTACCG CATGTTCTCC CGTAAAACTG TGCGGTGCGC   3840

GGTACACCAG CAGTACTACC TCATCCACCC GTGTCTTTCC GTCCAAAATC CATCCGTGGA   3900

GAAACGTATG CGCACGTGCG CGCGTCAGAG CCTGCGCACG AGAAAAAAAG GACGCAACAC   3960

GCTCAATGGA GCTGCTCCCA CTCGTGCGGA CAATACCTAA CGCGGCAGGA CTGAGCGCCG   4020

TGGCAATGGC GACGATGTCA TCGTCGAGCG CATACTCATG TGCGCGCATC AGCTACCGTT   4080

CCCTCACGGC CGTGGGCGCA GGTGCGTGAA AAGGCAAGCC GCCTGCAAGT CCCACACGGA   4140

GAAAAAGCAG CGGCACCACT CCCGTACTCA GGGAAGCTCC TATACCATAA CGCACAAAAC   4200

GCAACAGGCG TGCATACTCA GCGGGCGCAC CCGAGAACAG TGCATGAaGC GCAGAaGACA   4260

CCACACTACA GAACACACAC CCTACGCAAA AGCGCACCAC CTTCTGGGTG AAAGACCCCC   4320

CAGCAGCATG AAACCCCACG CCGcCTGCAG GTGCCCTGCG CGTAACCCAA CGATCAAAGA   4380

TAAAGAGGTG CCCTGCTCCT AACCCGCAGA ATGCACCACA CATCGACACA TCCGCCCCAC   4440

CGACAGCTAG CAAGAGTGCT ATGCACACCA mGGCAGAAAA GCGCAACCCC GCAAAAtGCG   4500

TCCTGAGACC TCCTGTATGC GCGTAAAGAA CACAACCCCA CGCCTGAGCG CGCGCCGCAT   4560

aCCGAGAATC AGCGCGCCAA AAAGTACTGC GTTTAACCAA CCAAGAAGCA CATCAATAGG   4620

ATAGTGCACG CCCAAATaCA CACGAGAAAG CCCAATGACT CCTACAAATA GCACGCCCGC   4680

```
CACGCCCGTC CATGCAGTGC GCGCCAGCGG GtaCGCTCCT GTCCATAAGA CGACGGGTTC    4740

TCCCGATCCC CTACCTGCCT ATCGGTTCCA CAACCTGGCG CACAAAAGCA GGGAACAGGA    4800

GTCTTCGCAG ACGGATAgcT ACGCGCGAGC AACACAAACA AAGCACTCGC CTGTGCATGC    4860

CCGGAGGtGT AGAAAAACCA TCGTGGAACA CAAGTTTCAC CGACGGGTCA CGCACAAAGG    4920

GCCGCGGGAC ACGCAACAGC CCCTTCAGGG CGTAATTGAG CCCCTCGCTA CATGCCAATG    4980

CGTAGGCAAT GGCTAAACCC TTTCGGTACT CTACGCACCA CAGCACCCAG AGCGAACACA    5040

GGGCGATACC CTTCCCTCCA AAGAAGGTAA AAAGAACAAC CGCGTGTGTT ATCACAGGGT    5100

GCGCAgCCTG cTGCACCGCG TGTATGACGG ACAAGTTCCA GAATATAAAT TCTTCCATGG    5160

TGTCCTATCC TCACTTTGAC ACGCGCGTCT GCATCGATCC GCGCTCCGTC CTGTGCGGCA    5220

GATCCTCAAG CCAATAAACT CTATCCGGAG CGCGCTCCTG CACAAAAATG gCGCTGTAGC    5280

GCCCTGaATG CTGCACACGT ACGCTGCCAG GAAAATACTG CGTGTGCAAA AACCACACAA    5340

GATCGCACAT ATCGTCGCGC GAATGgAAAc CGAGTAACGG TAGTAGAGCG TTGCATCGGT    5400

AAGCAGCGCT CTGCgCTGGa TCTGCTGTCT CATGCGGGTC AACTGCTGCG CGTACATTTC    5460

CCCGCAATCC CCCATACCAC GCGTGCGCCG CAGTCAGGAG ATTAAGCGCG CGCGCTTCAT    5520

CCATGTAGTA CAACTCGTAC ACACCTGACA CTGCAGGTAC CGCAGTAGAA ATGCGATACT    5580

TGTCCACCTG GGTGAGTGCC GACCACGTTA GCACGTACAG AACCTGCTGC GGGkTTCCCT    5640

CAGGAACCCC AACGGGCGcA GGTCTCAGTT GCTTAGTGAT CAACGGCTCC AAGGACTGCA    5700

TGTAACGCAA ACTCCGTGAC AATACAAAC TGGGAAAAgA GAGAAGCACC CGCCAAGACA    5760

CGCGCAGGCC GAAcGAAGGC GCCGATCAGA AtCGAACTGA TGCATAAAgG TTTTGCAGAC    5820

CTCTCCCTTA CCAaTTGGGc ACGGCGCCGA GGACCCCTCA GGCTAACAAA AAAAGACGCA    5880

ATCGTTCAAG GGTAAACCAA CCGATACTCC AGGCACGCTG TGACCTTGCG CAAAGGGGAT    5940

TACCATGGAA AAACAGTCAC CCGCACAAAC TATCTCGCTC TTCGTGCTCC TCGCGCTCAT    6000

GTTTGTACTC GTGTGCATGC TGTTCGTACC CTACtAACGG TGCTTCTCTG GTCGAGCATC    6060

CTTGCTATCC TGCTTTCACC GTGTTATCGC gCACTGTGTG CaAGAATAGA TATGCaTGCT    6120

TTTACGCGTA CTCGACATCT CGTTTCTCAC ATGAATGGAG AGGATGGATG TACCGCGGCG    6180

ATTACCCGAG CGACGCGCTT TCAAAAAAAG ATGCTCGCAG CGGTATTTTC ACTTGTGATT    6240

ACCCTTCTGG TGACCACTGT ATTTTTTTTC ATTGCAATTA GTTTGTTTGG ACAGGGAAAG    6300

CTCTTGTTTG ACAAACTTTC GCTCTTCTTC AGGGAATACG ATCTATTTGA AGGTGCAAAG    6360

CAACGGAGCT TTACCGCGCT TATTTTTAAA CTTTCCCGAG GAACGGTTGA TATCTCTACC    6420
```

```
CTCAATGTGG AGGAGCATCT GCTACGGTTC TTCGGCAAGC ATGTAGAATC GGTGTTTGTG    6480

TATACACAAA TTTTTGTCAA AAACATCGCT CGCGCAgCCC TTTCCACGTT GTTCTTTAGT    6540

TTTACCCTAT ACTTTTTCTT TCTCGATGGG GAACATTTGT CCTGTCTGCT CATCGCTGCA    6600

CTACCCTTGA GGAAgCGCGC AAgCGcACaG TTGTTAGAAA AATGCAAAGA GGCAACGCGT    6660

CATTTGTTCA AAGGTCTATT CTCCATtGCT TTTTATCAGA CCTGCGTTGC ATTTGTGTTC    6720

TACGGAATCT TCCGCGTGGA AGGACCGATG GCTTTAGCAA TGCTCACCTT CTTCGCCTCA    6780

TTCTTACCAC TGGTcGgcTG CGCCTGsGTG TGGCTCCCAG TGGGAATTAG CATTGGATTT    6840

ACGAGCGGGT GGATGCGCGG CACCCTTTTC TTGTTTGTCG CTGGAAGTTC AATCACTATC    6900

ATCGACAGTT TCTTGCGCCC GTTGTTGCTG CAAAATAAAA TGCGCATCCA TCCATTGCTT    6960

ATTTTTTTCT CTATGCTCGG TGGGGTGCAG ACGTTCGGCT TTAACGGTAT GGTGCTCGGT    7020

CCTATTTTGG TTATCCTGCT GTTCACGGTT ATCGACTTGA CGCACGACGG GGAGTCTCAC    7080

TACACGTCTA TTTTCCACGA CCCCCCTGCT GCAGGTGTGC ACGCGCAGTC GATACACAGA    7140

CAAGGAAAAA AATAGGGATA TCTTGCTGCT CGGCGCCCTT TTTATTACCA TGCGGCCCAT    7200

GACGCGCGCG TGTATATTCG ATCTTGATGG AACGCTAACG AATACGCTGG GGACCATTGC    7260

CTACTTCGTC AATATGCAGG CTGCCCaTTA CCATTTACCC CCAATTCCCT CTGAAAAGTT    7320

TGCGCTGTTT TTAGGAGATG GTTCGCGCGC ACTGATTCAG CGCGTGCTtG CTCATTACGG    7380

CGCTGCAGCT CAGACTATTT CTGAGGATGA ATTTTTACAG CGCTACTGCC TCGCGTATGA    7440

GGCAGACTTT CTCCAACGCT GTACTGTATA TCCGGGGGTT CCTGAGATGC TTGTGGAGTT    7500

GAAACGACGC CGCATAGAAC TCGCCATTCT CTCCAACAAG CCACATTCTA TCGCGCAGAA    7560

GGTAGCGTCT GCTTTTTTTG GGGACAATGT TTTCTCAGTG GTGCTTGGCC AACGCGAAGG    7620

CGTACCCGTA AAACCAGATC CTGCTGGGCT TTTTGAGATC CTGCGTACCC TAAACGTGGA    7680

GACGGCGGAG GCGCTTTTCG TCGGAGACAC CGCCGTGGAT ATACGCACCG cGTcCGCAGC    7740

GCAAGTGCGC AgCGTGGGaG TGCTCTGGGG CTTTCGAGAC GAGACGGAGC TATCCCAGGC    7800

GCAAGCCCAC GTGCTTATCA GGACGCCCGC CGAGTTACTC CAGCACCTTT CTTTCTAGAC    7860

TCGCGGGTAC AAACTCAGAC GGAGCGCACG ACGCTCCCGG ATCCCTGCAg GGCACGAGCC    7920

GCTACTTCTC TTCACGCCCA ACGCAgTTCG CCCGCAGGGT ATAGCGAAGT CCACGCAGCA    7980

TCAGTGCCAG GGCGCCATCC CCAGTGATGT TACACGCAgT CCCAAAACTG TCTTGCAAAG    8040

CAAATATCGC AATGAGCAAA CCGGTTCCTG TGGTATCAAA GTGCAACACA TCAAGCACCA    8100

GCCCGAGCGA CGCAAGCACC GTACCCCCTG GAACCCCCGG CGCACCTACG GCAAAAATGC    8160
```

```
CGAACAAACA GGAGAACAGC ACCATATCTG CAAGAGAGGG CATGGACCCG TACAACATCT    8220

GCGCTATCGT TAGACAAAAA AAGGTCTCCG TCAGAACAGA CCCGCACAGA TGTGTGGTTG    8280

CACCCAGCGG GATCGCAAAA TCCACAATTT CTGCAGGCAG TGCCCGTGAC TTGTGCGCAC    8340

ATTGTAACGA AACCGGCAGT GTTGCTGCAC TCGACATCGT GCCCAGCGCA GTCGCATACG    8400

CCGCTCCATA ATGACGAAtA CCTCGAACGG ATTTTTGCGT GACAGTATCC ACCCCACCAG    8460

GTACAACACG CACAGCCACA GGAGATGACC CACAATGACG ACCGCTACCA CTTTGGCAAA    8520

AAGCGGCAGc TGACGAGTTA AACTCCCGCT GTACGCAAGT TCTGCGAAGG TAGCCGCCAC    8580

AAAAAAGGGA AGCAGCGGCA CCAACACTCG GCTAATAGCT TCACCCATCA TGCGACGAAA    8640

TTCATACAGC ACCTGCTCCA CCGCACGTGC TTTTACCCAG AGGGCAGACA GCCCCACCAA    8700

GAGAGCAAAA GCAAGTGCAG TGACCACGGG CATAAGAGAC GGAATCTCAA GGGTAAAGAT    8760

AACCTTAGGG ATTGTACGCA AACCCTCCAC CGTGCGCGGG ATCCGAAGAT ACGGGATAAC    8820

AACACGCCCC ATCGCGGTGG CAAAAAGGGA GGCACCCACC GAAGAGAGAT AGGAAAGTAC    8880

CAGAAACGAG CCTAGCATCC TACCGGCACT CGCTTTCAGA CTCAGGACAG TAGGGGCAAT    8940

AAAACCAAAA ATAACTAGGG GAATAACAAA AAAAACAACC CCGCCGATAA GCGTTTTCCC    9000

CGTGTGGATA ATGGCCATGA CCGACTCATT AACGCACAGC CCGAGCGCAA CGCCACAGAC    9060

CATCCCCCCA CTGAGCTTTG CGAGCAGCCA AAACCCCGCA CTCCCGGCCA TAGCGTCCTC    9120

CTCCGCACAC GCGCCGGcAG TATACCAAAA AGACTATCCT CTGATAACAG GTCAGCGGTC    9180

TTTTTATGTC ATAGAACCAA CCTCGAAGGC GAGGCAAAAC AGATCGAACC CGCACCTCCC    9240

AAGAACTATG CAGGAAAGAC GCACCGACGG GTTGCATGCC GGGGCGGAGc GCACCCCTAT    9300

GCAACAACCC AGCTTACTCC CACTACGGTT CACTCAAAGA ATGTTCTCGA ACTCCTCCCT    9360

CACCGACCGC GGCCATACAC TGGCCTGAAC TTCACCAATA TGCTTCCTTT GTAGGAGTAG    9420

CATCGCCAAA CGGGATTGAC CGATACCACC TCCGATGGAT TGAGGAAGAC GACCATTGAT    9480

CAGATCCTGG TGCCAGyTGC ATGCCAGACT ATCCTCATCG CCaGTAAGAG cCAGCTGCGT    9540

GCGAAGCGCA CCCTCGTCCA CCCGTATCCC CATCGAAGAC ACTTCAAACG CACGCCCCAA    9600

CACTGGATTC CACACCAAAA TATCGCCGTT CAGGCCCTTG TATTCCCTTC GGAAGGCGTC    9660

GTCCAGTCAT CGTAATCTGG AGCGCGCACA TCGTGCGGCT TGCCGTCAGA AAGCACACCA    9720

CCGATCCCAA TCAGGAACAC CGCACCATGC TCTTTGCAAA TAGCATCCTC ACGCCCCTTG    9780

CTGTCTAAAT GCGGATAACG CCGCACCAGC TCCTCGCTCT GTACAAATAC AATATCCGCA    9840

GGCAAAAACG CCCGTAGgCC GAACCTCTCA CTTACCAACA CCTCCGACTC CCGAAGAGCA    9900
```

```
CCGTAGACCT TACGCACCGT GTCCTTCAGA TACGCAAGAT TTCTCGACCC TACCGGTACT          9960

ACCTTCTCCC AATCCCACTG ATCCACACAC ACAGAGCGCA CCTGATCCAA GAAATCTTCA         10020

TCCGGGCGGA GcGCGATCAT GTGTACAAAC AAACCCTCAT TATCCTGAAA GCCGTAGCGG         10080

GCAAGCGTGT GGCGnTTCCA CTTTGCTAAC GAGTGcACAA CCTCAAAGGC AGTACCCGGG         10140

ATCTGCTTcA CGGAGACGGA AACCGCCTTC TCCCGACCTG AAAGACCATC TTGGATCCCG         10200

TCACCCACCT GGCTCAGAAG AGGTCCCTGA ACTTCTATGA GTCCCAGGTG CTCCATCAGC         10260

TTTTGGGTAA ATGTGTGCTT GGCAAAGCTG ATCCCCTGCT GTTGCAAAAT AAATGATTTT         10320

TCCATAGTTA ACGCCAACCT TTTACCTTGT TGAGTAAACT GTGCACGCAT TATTTAATAG         10380

GGTGGCGGTG TAGTGCAATA CTCAAGTAAT CTGACAGCAG GGAGGTGGTG TGAAAAAACG         10440

AATGTGGCGC GCGGTGCGGA CCCTGCTTAT CATCTGTGCG GGGGGAACCG GAGCGCTGTG         10500

GGCGCATCCG CACGTTTTTA TCCGCACGAA AGTAACCTTT CAGTGGCAGA AGGGGGTGCT         10560

TCAACGCGCG CATATTACCT GGGAGTTTGA TCCGTTTTTC AGCGCCGATA TCATTAGCGG         10620

ATACGATACC AATAAAGACG GGCTGTTTGA CAAAAAAGAA ACACAGCAGG TGTTTGAAAA         10680

TGCCTTCATC CATACCAAAC ACTATTCTTT CTTTACCTTC ATCCGTTCCG GGGAGTCgCA         10740

TGCGCGACGT gCTCGCTCTC AAGCAGCACG TACAAGTCCC CAGTCAGTGC AGCATTTCTC         10800

GGTCAGTCAG AAAGACGGTA CGCTGTCTTA TCACTTCTCC ATTGACCTTT CTAGCTACCA         10860

GCACGCTAAG TCCGCACCCC CAGGAACCCG GCGAACACTG TATCTTGCAC TCTATGACCA         10920

CTCATTTTTC TGCGACTTTC GTTATGCAGA ACACGACACC GTACGCTTTG TGTGCGATAA         10980

GGCGCGCGTG CAGCCTTCCT ACGAAATTGT TGAAAACCGA ACCGCTCCTG TGTACTACGA         11040

CCCCTTCGAT AGCATAGAAA GCACTCCCCA ATACGAACAC TGGCGTCCCG GTCTGCATAC         11100

CTACTACCCA AAAGAGATTC TCCTGCGCTA CACTGCCCCC TAAGGTCCTT TTCCAAGGGG         11160

AGTTGAGAGC GTATGAAGAA AGTAGGGGTk cgCGTTCGCG CGTGTATCCT GTGCGCGCTT         11220

GCCGCGTGcG CCACAGGCGT CCTTGCTAAT CCTTTTTTTG GCGgcGCTCC CGCGCGCCCG         11280

CGgAGGCAGC GCACCCCGGA GCTTTTkcTG CGCAGATACG CGCTCGTCCA TCAACGCCTC         11340

GGTGCCGCCA TAGTACAGTG GAGCAAAACC CATTCAACAC GCGCGTGGTG GATTACTGTA         11400

ATGCTCTCCT TTGCGTATGG CGTTCTGCAC GCCTTAGGAC CAGGACACAG AAAGGCAGCG         11460

CTTTTTTCTT TCTACCTGGG GAGGAACGCA CCTGTGTGGG AACcTGCGCT CACTGCAGCG         11520

TTACTTGCGG CGTTGCATGG CGCAgcTTtC CCTGCTCTTG CTTTCTGCAT TTAGAGGTGT         11580

TTCCGGCGCA ATCGGTGCAC ACAGTGCACG CACAATGTGG TACATGGAGG TGGGTTCCTA         11640
```

```
CGGATTGCTC ACCTTCTTAG CGCTTTTCTC TCTCGTGCAT GAGCTGATGC ACCTTTTCCc   11700

TTCGGGCGGG CGCTATTTCT CCTGCGGTTG CAGCGCGCAC ACTGCCGTGT GTATGCGGAC   11760

AGGAACAGTC GCCCACATGC AGTGGGGTAC TATGCTCTTG AGCGGTTTAT TTATTTGCCC   11820

TGCTGCGTTG TTTGTGATGA TTCTGGTGCT CAGCTTAGAT GCAGTTGGAC TTGGCGTCGC   11880

AGCGGTGCTC AGTATTTCAG CGGGGTTAGC ACTCCCCCTG ATGGCTGTCG GTTATTTGGC   11940

CTGGGCGAGC CGGGCAGGTA TTTTTTATCG CATGCAGAAG AACACTCGTC ATGCACAAGC   12000

GGTGCTCTCT GTCGTGAGCA TTACCTCATA CGGAATTATG CTCATCGTCT GTACTTCAGC   12060

GCTCGTAGCT TCACTCGGTT GAAAGGAGAA TGTACCTCCG CTATCTAGGT GACACTGCCT   12120

GGATAAAACC ATATACCTAA CACGTGGTGA ACGGAAGTAC GCAGTATCTT GCACACGTCG   12180

GTGAGCTCAG CTTAAAGAAG GGGAACCGTA GACAGTTTGA AGTGCAGCTT GAGCGCAACC   12240

TCACGCTCAT GCTACGAAGC ATAAACCCTC ACGTTACTGT CCGCGCAGGC AGGCTGTATC   12300

TGTCAGTCCC GGCCTCCTTT GAAGCACAGA CCACCGCTGA GCAAGCCCTC TCGTACCTGC   12360

TGGGAATTAC CGGTTGGGCT GCTGCTACGG CGTGCCCCAA AACTATGGAA GCGATCACAC   12420

GGTGTGCACA TGCTGAGGCG ACGCTCgcTG CGCGCGAAGG AAAGCGAACA TTCAGAATAG   12480

AGGCgCGGCG CgcGGaACAA ACGCTTCTGC CGTACCTCGA GTGAGATTGC ACGGGAAGTC   12540

GGCGCGGTTA TCCACCAATC AGGCGCTTTG TCCGTGGATC TCCATCATCC TGACGTGGTC   12600

ATTTTCATAG AAGTGCGCGA GCGCGAAgCC TTTCTGTATG GTGCCCGACG TCGCGGCCTG   12660

CGTGGTTTAC CCTGTGGCGT CTCAGGACGC GGGCTACTCC TGTTATCCGG CGGCATTGAC   12720

TCCCCGGTAG CCGGGTACCG AATGCTTTCT CGTGGCATGC ACATTGACTG TCTGTATTTC   12780

CACTCTTATC CCTACACCCC TCCTGAAGCA CAGAAAAAGG TTGAAGACCT GGCAAAGGTA   12840

TTGGCGCGCT ATGGACTTAG TACCACGCTG ACAGTCGTAT CGTTGACAGA CATTCAAAAA   12900

CAGCTCCAAA CACACGCCCC TGCCCCTTCC CTCACACTGT TGCTTCGTAT GTGCATGATG   12960

CGCATTGCAG AGCACGTAGC GCGGGAACAG CGCGCACGTT GCCTTATCAC TGGAGAAAGC   13020

CTTGCACAGG TAGCAAGTCA GACGCTTGAG AACtAACGGT GACCAGCGCG TGCACGCATC   13080

TGCCGATATT CCGCCCGCTC ATTGGTGCAG ATAAAGAAGA TATTATCCGC ACCGCCACAG   13140

AAATCGGTAC GTACGCCATT TCTATCCGTC CGTACGAGGA CTGCTGCACA CTCTTCGCAC   13200

CAAAACACCC AGTGcTTCGC CCAGAGGTAG AAGAAATGCA AAAACAATAC CAATCTCTGA   13260

TGCTCGGTCC ACTGTTAGAA GACGCGTTCC GGACGCGCAA ACGCACGCGC ATATACGGAA   13320

ACTATGGGGT ACAGGAGTCA GGCGAATGAG TACCGCTTAT CTTACGCGGC AGCACCGTCC   13380
```

```
GCCCCTTCTT TAtGACGCGT TTATGCTAAC CAtCAAGgAT TATTcCACCA GCGCTTCGGC    13440

GGTAAATTCC AGCGCCTGTG CCACCTGcGT ACCAGGGCCT GACTCTCCGA AACgGTCGAG    13500

CACAAGGCAT TTTTCCCGCT TTGCCCACGC TCCCCAGCCT TGATACACCC CTGCCTCAGC    13560

CACTACAACG CGTGCTCCTC CCTGTATGCG CCGCTGCACC TCGTCCCcTG CTGCCTCAAA    13620

ACGCTCCTTG CACAGTACAG ACACCACAcG CACACGTCTT TTACTcAGTG CGCGGCACGC    13680

AACGCCAAAT CCACCTCAGA GCCACTTGCC AAGACAGTCA GCTCAGGCGT AGCACCCCCT    13740

TCGCGCACTA CATAGGCCCC CGACTCCTCC ACCGTAGAGC GCCACGAACT GTCACTTTTC    13800

TCAAAAACCG GCACGTTCTG CCGACTCAAA ACGATACACA CAGGACCAcT GCGGTGCAGC    13860

AACGCTATTT TCCAAGCTTC AAACGTTTCT TCTGCGTCAG CAGGGCGCAG AACAAGCACG    13920

TTGGGAATCG CACGCAGCGC AGCGAGCGTC TCCACCGGTT GGTGCGTCGG CCCATCTTCT    13980

CCTACAAAAA TAGAGTCATG TGTTAAAACG AAAACAGAAG GGATGCGCAT GAGCGCCGCA    14040

AGACGGAGCG CAGGGCGAAA GTAGTCTGAA AAAACCATAA ACGTAGCGCC AAACGCACGC    14100

AAACCGCCGT GCAACTGCAT TCCGTTCACA ATGGCTGCCA TGGCAAACTC GCGCACACCA    14160

AAATAACAGT AGCCGCCTGC ACGATGCTCT GCAGAAAATG GTCTTAACGA AGAGACCGCT    14220

ACCGCATTCG GCCCGCGTAA ATCTGCAGAG CCACCTACCA GATTCGGTAG CACAGAGCAG    14280

AGCGCGTCGA GCACCTTTCC AGAAGCAGTC CGAGTAGCAA GTGACGAACC CTTCTCAAAA    14340

TGGGGACAGA CAACACGAGC TAGCTGCGAA GTACTTAnCC CTCCGGGAAC AAAAGCAGCG    14400

TCCCAGTCAG CACGTTTTTC AGGATAtGCG TGCTCCATGC TTCAAAGAGC TCATTCCACG    14460

AGTCCTCGAC ATGCGCACAT TCACACTTTC GTTTCTGGAG AACAGCGGTA AGCTCAGGCG    14520

CTACAAAAAA AGAGCACGCA GGATCAAGTC CCAATGCCTT TTTTGCCTCT CTCACCCCCG    14580

CTTCCCCAAG CGGGGCGCCG TGGGCACGCG CGCTCCCTTC AACGGTAGGC GCACCCTTTC    14640

CAATAATCGA ACGCaGGATA ATGAGAGAAG GCCGATCGTC ACGCTTTGCA CACGCAgTGA    14700

GATCCATAAT ATCCGTATAC GAATACATAG AACCGCGCAg CACCTGCCAG CCATACGCTT    14760

CGTAGCGCTT AGCCACATCC TCGnTAAAGT CAGATCGGTA GATnCGTCTA TGCTGATGTG    14820

GT                                                                 14822
```

(2) INFORMATION FOR SEQ ID NO: 45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16710 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

```
TGCATCnGAG ATACACAAAC nGTTTCTGCC CCTTAAAGCT TCAACTGCAA gTACTTTTAG      60

CTGAATGACG TTGGCGCATT CTCGCGCGAG TGTTTCTTCC GTGCGCGGCG TGATAGCTTC     120

CGGTCTGAGG CCAAAGACTA CCTGTGTGCC AATATAGTTT TTTAACAAAA AGGACGCGGG     180

GATATCCGGT CGAAGAACAA AAAGACCTGC ATCAATTTTT ATCCCATGCT CATCTTTAAC     240

AATCGTAACA GGGAAACAAT TCATAGGAGG AGAACCAATG AATTGTGCAA CGAACGTGTT     300

CGCAGGATGC TGGTAGATAT GGAGAGGAGA ACCAATCTGT TGTACGCACC CGTCTTTCAT     360

GATGACAATC TTATCTGCCA TTGTCATCGC TTCCATTTGA TCATGGGTGA CGTAAATCAT     420

CGTGGCCTTT AGGCGCTTGT GAAGAAGAGA GATTTCGGAT CGCATTTGCA CGCGCAACTT     480

TGCATCTAAA TTTGACAATG GCTCATCAAA GAGAAAAACC TTAGGATTTC TGACAATGGC     540

ACGTCCAACT GCAACGCGTT GTCGCTGTCC CCCCGAAAGT GCTTTGGGTT TTCGGGCAAG     600

CAGTGGTTCG ATATCAAGAA CACGCGCTGC TTCGTGGACA CGGCGGATGA TTTCTTGCTG     660

AGGGATTTTA CGGATTCTAA GGCCGAACGC CATGTTGTCA AAAACGTTCA TGTGTGGGTA     720

GAGCGCGTAg TTTTGAAAGA CCATCGCGAT ATTGCGATCT TTTGGTGTAA CGTGATTCAT     780

GTGCTCACCG TCAATGTAGA GGTCACCTGA GCAGATATCT TCAAGCCCTG CAATGATACG     840

TAtGCAGTTG ACTTGCCGCA TCCAGATGGT CCGATGAACA CCACGAACTC TCCACTTTCT     900

GCGGTAATAG TTACGTCTTT TACTGCATGG ACGCATCCGT GATACGTCTT ACAGATATGC     960

TTGAGTTCAA CCTTTGCCAT AGCGTTTACA TTCCTTTTGA AACACGGGTG CGCAACACAC    1020

TACTTTCCTT ACGCAAACGG GAGGTGGTGT TGTCATGTTA CGCGCTCTGC ATGTGGTGCA    1080

AGCGGTCGTC TAGATATGCG ACAAGCGCTT CGGTAAACAA ATCTTGATTT ATCAATTCGT    1140

GTGCAAAGAA ACGGAAGCCG ATGGCACCGG CGTGTCCACC GCCGTTTTGA ATAGCAAAAT    1200

GAGATAGCAG TAAACGCAAA TCCAACTGTT TCACCCGCGC GGCAGTGCGC ACGCGCCACT    1260

GCGTGACATA AAATTGCTGT GTTGGATCTG GATACGCGGA GATACCGATA CCTGAGATAC    1320

TCTCTGCAAT GTTGCTGGTC GCCATTTTTA TCAACGAAAC GAAATGTTCT GCATTGCAAC    1380

ACTCGTGCAA CGCCTGTGTT TGTGTCTGAT TGAGAAGGAG CAGGGAAATA CTGCCGCGGT    1440

GTTGCACGTT TTTCACCATT GTTTGGTAAA CCGCATGTTC TTCAGCGGAC AGAGACTCCA    1500

GGGTATGGAG GATTTGTTTG GTGGATGCTA TGTTCCCGGT AATCGGTTTT GTTTTTTCCC    1560

TGAGCATAGT GTTCAGCCTC TGGGTGAAGT AGGTGTAGAG CGCGCGGTCT TTCCGACTTA    1620
```

```
TCAAATAAGC ACCTGTCTTT GCATCCCCTA TCATACCGGT GAGTATTGAA AGCACAACAT    1680

TGCGTGAATA CAAATTGCGT ATCCCAAGCG CAGTgCGCTG CGCGTgGTTA CGCGCGAGTT    1740

TGTAACAAAG ATAAGCGATG ATTTCGCACG TGCTAGAGGC GCGTGCGATA AGGCTGTAGC    1800

CTGGATCACC GCAGCAGGCA GCGTTTGCAA AAAGGTGATG GTCAAGTTCA ATTTTACGGA    1860

TAGTCGAATC TGAAAGGAGT AGTCGGCACG ACGGAGGCGC GTAGATCATA CCGGGATTTG    1920

GGGTGTCTAG AATGACCAGG GCATCCGGCA TACGGGGGAC AGTTTGTGTA TCGAGATGCA    1980

CGGCAATTCC GTTATAGAGA CAAATATCAA TGAGGAATGA AATCTGCACA CGAATGGGAC    2040

CTTGACAACA AATTTCCACG CGTTTATTGC ATCGCGTGAG CAGGAGTGCG AAGGCTACTA    2100

ACGATGCGAT ACAGTCTTCA TCAGGATGTT CATGTCCGAG CAAAAGAAAG GAGCCGTGTA    2160

TCGCAATCTC CTGCAGAATA TTTCGGACGA CGGCATTTTT CGCTGCAATG GAGAGATCTC    2220

GTTTTGGCGA CGGAGAATAG GTGTAGTCAC cACTAGCGCG CCCATGACGG CGAGTATACT    2280

ATAGAGTGTC CGCTGGTGTG CGGCAAATGA GACCGTTCTC CCAATTATTT TCTGGGTAAT    2340

GTTGCTCCCT TGTTACTCGT GTGAATTTGC AGTACCGTTG GGCCAGAGGC AGCTTGGCGT    2400

GCGCAGGAGT GATTGTGTGT GCGCCAGACG TGACGGAGCC GAGGTTGTAT TGAGTTTTCC    2460

TTTTGCATGA TGACGTTTCC CGAGGCTGCA TCGTTTTTGC CGTTCGTGCG GCGCGCAAAG    2520

GCGTGCCTGC GCGTCTTGCG TATATGAACG GACTGCAGCC CTGCCTTGTG CAGGGTGTGT    2580

CTTGCGATTG CGTTGGGGCA TGCTGCAAGA AGCGTACACA GGGGGGCATC CCCCACGAGG    2640

GTATGGAGAG GAAATGAAAA TTATCATCAG TGCTTCTGTG CAAATTATTC TTGATCAGGC    2700

CTTTGATTTA GCGCGTAAGc GGCGTCACGA GTACATCACC GCAGAGCATG TTCTTTTTTC    2760

TGCCCTTGCG CATCCTGCTG CGTTGGAAAT TATCAATCTC TGTAGCGCGG ATATCGCGCT    2820

CATCCATAGT AATCTGTCAG AGTTTTTAAA AACACAAGTT CCCGTTGACT TAAGTCATAC    2880

TCCTTCTCAA TCACTGGGTT TTCAGCATCT GCTTAAGCGT GCAGTCTTGT ATTGCGAGGC    2940

GAATAAAAAA AGTGCTCTTG AAGTCGGAGA TCTGTTGGTA AGTCTCCTCC AAGCGGAGAC    3000

AAACTATGCT TCGTACTACA TGCGTATGTC GGGTATGAGT ACGGCGCGCT TGATTGAAGT    3060

AATAGCTCGT GTCAATGGCA TCCGACACGG GGATAAGAAT GTGTCGATGG GGTCGAACGC    3120

GCAAGAAAGG TATTCAGAGT CAGACGACGT TGGCGAATCT GCAGGGCACG GTCCTCCGCT    3180

CGATGGAACA GAGGGGGATG GTAACACAGC GGACGTACAT GTGCACTATG AACACTGCGC    3240

GCATAAACGC ACGGATGCAG ATACGCATCG GTATACGGTG CTGGAAAAGT ATACGGTGAA    3300

TCTCACCGAA CGTGCTCGTC GGGGAGAGCT TGCTCCGCTC ATTGGGCGTA CGCAGGAAAT    3360
```

```
TGAGCGGACG ATTCAGATTT TGTGCCGGAG ACAGAAGAAT AATCCGGTAC ATGTGGGTGA   3420

AGCTGGTGTG GGAAAGACGG CAATTACTGA GGGGCTTGCG CAACGTATCG TGCGGTGCGA   3480

TGTGCCAGAG GCGTTAGAGG GAGTAGAGAT TTTTAGCCTT GATATGACAA GCCTGTTAGC   3540

AGGTACAAAG TTTCGAGGGG ATTTTGAAGA GCGGCTCAAG CGTCTTGCAG AAGAGTTGGA   3600

AAAGAAAACA CAAGCAATTC TTTTTATTGA TGAAATTCAT ACGGTAGTCG GTACTGGCTC   3660

AGGCGGTTCG GGTGGTTTGG ATGCGTCTAA CTTACTCAAA CCGCTGCTTT CTTCAGGAAA   3720

GATTCGCTGT ATTGGTTCTA CCACGTATGA GGAATACACC AAACATTTTC GCAAAGATCA   3780

GGCGTTAgcA CGGcGTTTTC AAAAAATTGA TATTGAAGAG CCTTCTGAGG AGGAAACCCT   3840

CCGAATTTTG GAAGGGATTC GCACGCTTTA CGAAGACTTT CATGCAGTGC ATTACAGTGA   3900

TGAAGCATTA GCTGCTGCGG TGAGACTTTC GGTGCAATAC ATCCAAGGGA GACATCTGCC   3960

GGATAAGGCG ATTGATATTA TCGACGAAGC AGGCGCGTGT GCAAAGCTAT CCCGGGGAAA   4020

GCACGGAACA GAGGGAGTGT GTTCAGTAAT TGGGGAGTCG GATATAGACG AAATTGTGGC   4080

AAAAATTGCG AAAATCCCTA AGCAGCGGGT ATCTGCAAGT GAAATAGAAA AGTTGCGTAA   4140

CTTTGAGCGC AGTATTTCAG AAAAAATTTT TGGACAAGGC GAGGCAATTG ACTTAGTCAC   4200

TCGTACgCTG AAGCGCGCGC GGGTGGGATT GCGCGTAAAG CATAAACCTA TAGCAAACTT   4260

GCTTTTTGTG GGGGCTACCG GTGTGGGAAA AACAGAGCTT GCGCGGACGC TTGCCCAGGA   4320

ACTAGGGATT GTGCTGCATC GTTTTGACAT GAGTGAGTAT CAGGAAAAGC ACACGGTGAG   4380

TCGGTTGATC GGCTCACCGC CCGGTTATGT TGGGTTTGAA GAGGGGGGAT TGCTCACCGA   4440

CGCGGTAAGG AAACAACCGC ATGCGGTGCT CCTTTTGGAC GAAATAGAAA AAGCTCACCC   4500

GGACATTTTT AATGTCCTGC TCCAGGTTAT GGATTACGCA ACGCTCACTG ACAACCAAGG   4560

CAGAAAAGCG GATTTTCGCA ATGTTATTTT GATAATGACA AGTAATGCGG GTGCCCGGAA   4620

CATGGGTGTT TCTCTCATCG GTTTTCACAA GGGGCAGGTG GGTACTGCAG TTATCGACGA   4680

AGCAGTAGAA CGTATTTTCT CTCCAGAATT TCGGAATCGG CTGGACGCAg TTATTCGTTT   4740

TGATGCGTTG TCCTTGGAAA CGATGGAACG CATCGCCCGC AAGGAGCTTG CCCTGGTGTG   4800

TGAGCAACTG CAGAAAAAAC ACATTCGTTT TGATATTACC GATGATGCAC TCGCGTTGCT   4860

CGCTGAGCGT AGTCACTCAG GGGGAAGTGG TGCGCGTAAT GTTGCACGCT TGGTAGAGCA   4920

AGAAGTTGCA AATGTGCTTG CAGATCTTAT GCTTTTTGGA GGAGTCGCTG AGGGGGATGC   4980

GTTGCGGTGC ACGGTAAAAG ATCGGCATGC TCAATGCAAT TTTCTCCGCA TCGAGTGCGT   5040

GCAGTCTTCG TATTCGGGGA GTATCCAAGA CGCGCTGGGG TGATGATGCG TGGCACGGTA   5100
```

```
ACGGTGTATC CGTGTGGTAG GCAGTGCACG TGCACTGAGT TATTCAAACG CGTTTCTCTC    5160

TTATCTTCCG CCAAAGCTCT ACACTCCATA CCGCCACCTC GTCTATGATC CTGCGTGCGT    5220

TCCCACCGGC AGTGAGAGTG GTGCCTGAAA CGAGCGGAAC gCaGcGAGCA CCATTCCATT    5280

TGATTCGTCA AGAACATGCG CAAAACCGAT AACGTCCGTT TCTTGTAGCG GTGCGCGCAA    5340

AAGGGGAGGG AGTGCGAACG TTATGCTGAT GCGCGTACCG GGCGTATTCA GAACTGGGCA    5400

CGATGTACAC GAGGGGTGGA GGATTGGTCG CAGTGCCCCA GGGCGCTTAC TGCCCAGGAC    5460

GGGAATAGCA CAGGGAAGCG CATCACTGAT AGCCGTGCGT ACATCGGTGC ACTGGAAATG    5520

GGTGAACGCC CAGTATAGAA GGGTAGCGCC GTCACGCGCG CGGATGCgCT TTCCTTCGGG    5580

AATGGAATTC CCTGCGCCGC CTAAGATAAC TGCAATGATG CGGGTGTCTC CGCGCAGACT    5640

GCTGAGCGCG ACGTTAAAAC CGGATTCGCG GATATAGCCA GTTTTTAATC CGTCGCAGCC    5700

GGGCACTGCT GTGGAAACGT TTCCGCGCGC TGCGGTTGCT GCAAGCAGGG TGTTGGTTGC    5760

AGGGAAACGT GTGTGTGGGG GTATTCTCGA AGTAGCGGGA GAGTTACTAC GCTGAAAGTA    5820

gCTGCGCGCA TGAAAGCGTG CAAGGTTTTC AGGCCATCGG CGCACATACT GGCAACAAAA    5880

GAGCACAAAG TCACGCGCAg TAGTTACATT GTGTTCGCTC AGACCGCTTG GTTCCACAAA    5940

GCGTGTGCGC GTAAGACCCC ATTTCTGTAC AAGCGTGTTC ATGCGCGTGC AGAACGCCGG    6000

TATACTGCCT GCAACTGCAT AAGCGAGGGT GTAGGCTGCA TCGTTCCCCG AAGCGATGTT    6060

CATACCTGCG AGTAAGTCGT GTACGCTGAT GTACTCACCT GTACGTAAAA ATATAAGCGA    6120

GCTGCCCGGT GCAAAGGCCT GCGCGCTACC TGCAAGCGGT ACGCGTATAC GCTGTTGCCA    6180

GTGGAGTTCT CCACGCTCGA GTGCTTCCAT GACCACTGCA CAGGTAACCA GTTTTGCCAA    6240

GGACGCCGGG GGGAGGGGTA GGTCTGCGCA GAAGGAGGCA AGGAGTGTGC CGCTTCCTCC    6300

TTCGGCGATG GCGTACGCGC GGGCACTGAT GGGGGGTGGG gTGGATCCTG CAGATAGATT    6360

GGTAAATTGA AGCGTACGGA TAGGGTGAGC AGAAAAGGGA TTACGGGACG ACGGAGAAAA    6420

ACGCAkTgTT GGGGAGAAAA GAAAGGAAGG GTGGAGTACt CkCTGCGCGT GCCACTGCAG    6480

CGCCCCTGCC CCGACTACGC ACGCACCTAA CGCATACAGC GCGCGCTGCC CACGCCGCCG    6540

TGCCACGGCG CGCAGGGAGC GCGATGAGTG .GTCTTTCAAA CGGGCAGTAC ACGCGTGTAC    6600

CGGGTGTCCC GGCACACCGT ATTAGGGACA GAGCGTACGG CACGGGCCGA CCGCACATGC    6660

GTCACCCCTT GAAAAGCACG CCGCGACACG TCCTGTGTAT CACAGAAAAC ACCACACGGG    6720

TCATCATACA GCCTCCCGGC AGAGCATGTT GTTTGCATTA CTTTAGTATA GCAGAATGCG    6780

AAGTGTGCAG CGAAGGATTC ATCAATCCTG TTGCGTTCTC TTCTTTTTTG TGAGGCATAT    6840
```

ATTGCACCGG ATGCTCCTTG CGTCTGCCTG CCCTTTTGGG CAACACTTGT TGCCAGAAAC   6900

CTGTCTTCAC GAAgTCTTGT TTAAATCAGA GTAGGAAACG CTATGACGCG AAAATTAATC   6960

ACCGCCGCAC TCCCCTATGT GAACAACGTT CCACATTTGG GAAATCTTAT CCAGGGTCTT   .7020

TCTGCAGACG TTTTCGCACG TTTCTGTCGG ATGCGCGGCT ATCACACGTG TTTTGTATGT   7080

GGTACCGACG AATACGGCAC GGCAAGCGAA ACCCGTGCGG CAGAACAAGG TCTCAGTCCT   7140

GCACAATTGT GTGCGCACTA CCATGCACTG CATCGCGACA TCTATCAGTG GTTTGATCTG   7200

TCCTTCGATT ATTTTGGGCG CACTACAAGC GATGCGCATA CTGAGcTTAC GCAAGCGTTG   7260

TTTCGTCATT TGGATGCGCG GGGTTTTATC AGTGAACATG AAAGTGCGCA GgCGTACTGT   7320

CTGCACTGTG CACGGTTTCT TGCTGATCGC TATTTGCGCG GTACCTGTCC CCATTGCCGT   7380

AATGCTGAGG CGCGTGCTGA CCAGTGCGAG CACTGTGGAG TGCTCCTTGA GCCGGAAACG   7440

CTCCTGAATG CGCGCTGTGT GAGCTGTGGC ACGGCGCCGG AGTTTCGCCC TACGCGTCAT   7500

TTGTATTTAA ATTTGCCTGC ACTGGAAAAA GCCTACCGCT CGTGGTTTTG CACCACGAAT   7560

CATCTGTGGA CTAAAAACgC GGTGCgTATG ACTGAAGGTT GGCTACGTAC GGGATTGCAG   7620

GAGCGTGCGA TCACGCGCGA TCTGCGCTGG GGGGTGCCAG TTCCCAAAGC AGGATTTGAG   7680

CAGAAGGTAT TTTATGTGTG GTTCGATGCG CCAGTCGGTT ACATTTCCAT TACTAAGTGC   7740

GGCACAGAGG CAGCTTCCTC GCAAGAAGGG GGGGGGACCG ACGATGGCGT GAAAGAAAAA   7800

TGGCAGTCTT GGTGGCTTGA TCAGCAGGAT GTGGAGTTGG TCCAGTTTGT GGGGAAGGAC   7860

AATATTCCCT TTCATACGCT GTTTTTCCCC TGCATGCTCA TCGGTTCCGG GCAGCGGTGG   7920

ACGaTGcTTA CGCGTCTTTC TGCGACGGAg TATTTGAATT ACGAAGGGGG aAGTTTTCTA   7980

AGTCTTTAGG GGTGGGCGTT TTTGGTTCGG ATGCAAAAGA ATCGGGCATT CCCTCAGATC   8040

TGTGGCGTTT TTATCTCCTG TACCATAGAC CGGAAAAAAG CGATGCGCAC TTTACCTGGC   8100

ATGAGTTTCA GGAGCGTGTA AACAGTGAGT TGATTGGTAA TCTGTGTAAT CTGGTCAATC   8160

GTACGCTCAC CTTTGTGGCG CGTACGTACG GGGGCGTGGT CCCTGCGCAA GATGGAGCGC   8220

GCAgCACCCG TGCGCAGGTG ATGGAAGAAA CGCTTGCGCT CCGCGAAGrt GCGGgAATAC   8280

TGCAAAGCGC ATGACAGATT TAATGGAGCA GGTACAGTTG CGAGAAGCGT TTAGAGAAGT   8340

GTTTGCGCTC TCAGCGCGTG CGAATAAGGC GTTGCAGGAT GGTGCACCGT GGAAAACGCG   8400

GGCGCAGGAC CGTGAACGTG CAGACGCCTT GATGCGTGAG TTATGCTATG TGATTCGGGA   8460

TGTGCTGATT TTAGCGCATC CTTTTTTGCC GTGGTACACG CAGCAAGCGG CCCGATTTTT   8520

.GGGTGTTCAG TTGTCTTCCT GTGCACCAGA GGGGGGAGGA GCTGTGTGTG CTGCGAAGAA   8580

AGACGCGGAT ACGGCGCAng AACnACAGTG CAACCGACCC TCCGATGGTC AGACGTGGGA   8640

GAACGCAAGG GTTTAACGCA gGTGCATCCG CCGGTGATTT TATTCCGTCC GTTGGAGACG   8700

GAAACTATTG CTGCGTATCG TGCCCGCTAT GCTGGAACAc CAGGGATGGG GCAGGAGTGA   8760

GCGTACCGCG CACTGCACAG ATGCCCACGG GAATGAATAA GAAAGAGACA GACGCTCAAC   8820

AAAAGAAGGA GGAGCGTGAA ATGCCCCCTC CCTCAGATAC TGCACGGTTA TCTGCATTTT   8880

TTTCTGAGCG CGTTGTACTG AAAGTAGCAC GAGTGTTGCA GGTGGAGCGT CATCCGAATG   8940

CGGATATGCT TTTTGTTGAA ACATTAGATG ATGGCTCTGG CGTTGAGCGC GTTATTGTTT   9000

CTGGTCTTGT GCCTTATATG GCTGCAGATG CGTTGCGTGG TGCGCACGTG CTTATTGTGG   9060

ATAATCTGCA GCCGCGCTaC TGCGTGGGGT ACGGTCTTGC GGCATGCTGT TGGCCGCAGA   9120

GTATGTAGAT GCGCAGGgCA CAAAGGCAAT TGAATTGGTG CAGGCGCCAT GGGCTCTGCC   9180

CGGTGAACGC GCAACACTTG CGAGTGCGCC GCCGGTCATT ACACCGCACG GGTCTGCCGT   9240

TATCGATGCG GACGCTTTTT TTTCTGTGCC TATTCGTGTG GTAAATTATG CAGTAGAAGT   9300

TGCAGGTGAG CCGCTCATGG TTGGAGGAAG GCCACTGGTA ATGCAGCGAG TGAAAGAGGG   9360

AACTGTCGGC TAGGAATATT CACAGAGCAT TTGGTTTTCC GTGTCGGATA GGGGGAGCGC   9420

AgcATGAACG TGGGATTTTT GGGTTTTGGA GCAATGGGAC GGGCGCTGGC AGAAGGGTTG   9480

GTGCACGCAG GAGCGCTGCA AGCGGCTCAA GTGTACGCCT GTGCGTTAAA TCAGGAAAAG   9540

TTGCGTGCGC AGTGTACATC TTTGGGCATA GGTGCCTGCG CGTCAGTTCA GGAACTGGTA   9600

CAGAAAAGTG AATGGATTTT TCTTGCAGTC AAACCATCTC AAATCAGCAC GGTACTGCGC   9660

GATCGCCAAT CCTTTCAGGG AAAAGTGCTT ATTTCCCTTG CGGCGGGTAT GTCTTGCGCT   9720

GCATACGAGG CATTGTTTGC CGCGGACCCT CATCAGGGTA TCCGTCACCT GTCACTTTTG   9780

CCGAACTTAC CTTGTCAGGT GGCGCGGGGG GTGATCATTG CAGAAGCGCG CCACACCCTG   9840

CACCACGATG AgCACGCTGC GCTTTTAGCA GTGCTGCGCA CAGTTGCACA GGTAGAGGTG   9900

GTGGACACCG CGTACTTTGC GATCGCAGGG GTGATTGCAG GCTGTGCTCC GGCGTTTGCC   9960

GCGCAGTTTA TAGAAGCGCT CGCTGACGCA GGGtGCGCTA TGGCCTGGCG CGCGATCAAG   10020

CGTACCGGCT TGCGGCACAC ATGCTTGAAG GGACTGCAGC GCTCATACAG CACAGTGGTG   10080

TACATCCTGC ACAACTTAAA GATCGCGTGT GCTCTCCTGC AGGGAGTACT ATTCGCGGGG   10140

TGCTTGCGTT AGAGGAGCAG GGATTGCGCC GTGCAGTTAT ACACGCGGTG CgCGCTGCGC   10200

TCAGTTCTTC CTAAGGGGTG GGCAGGGTGC ATTGCTTGTT TTTTTGACT GCTGACAGTA   10260

CAGTTGCACC CTTGTGAAAA GTTCGTGCGT ATATTGGCGG ATCGGGGTTC TCGTTTGTAT   10320

```
TCTGTGTGGA GTGGGGAGCT GTGGCGgTCG TGCGCGCGTG CgcGAGTATT CGCGTGCGGA   10380

gcTTGTTATC GGTACGCTCT GTCGCGTGCG CGTGTACTCT AAGCGACCTG CTGCTGAAGT   10440

GCACGCGGCG CTTGAGGAGG TGTTCACGCT GCTACAACAA CAGGAGATGG TGCTGAGTGC   10500

TAACCGTGAT GACTCTGCGC TTGCTGCCCT AAACGCTCAG GCAGGTTCGG CACCGGTTGT   10560

TGTTGACAGG TCGCTGTATG CGTTGCTTGA GCGTGCGCTT TTTTTTGCAG AAAAGAGTGG   10620

GGGTGCGTTT AACCCCGCAC TAGGTGCGgT AGTCAAGCTT TGGAATATTG GCTTTGACCG   10680

TGCTGCTGTC CCTGACCCCG ACGCGCTCAA GGAGGCGCTG ACACGTTGTG ATTTTCGTCA   10740

GGTGCACCTG CGCGCTGGGG TATCGGTGGG CGCGCCACAC ACGGTACAGT TGGCACAAGC   10800

GGGCATGCAG TTGGATTTGG GCGCCATTGC TAAAGGATTC CTTGCGGACA AGATTGTACA   10860

ACTGCTCACT GCGCATGCTT TGGATTCAGC GCTCGTTGAT CTGGGAGGAA ATATTTTTGC   10920

CCTTGGTCTT.AAGTATGGAG ATGTGCGCTC AGCAGCcGCG CAGCGGTTGG AATGGAACGT   10980

GGGTATTCGC GATCCGCACG GCACGGGGCA GAAGCCTGCA CTGGTGGTGT CGGTGCGCGA   11040

TTGCTCGGTG GTGACTTCTG GTGCGTACGA GCGTTTCTTT GAGCGTGACG GGGTACGCTA   11100

CCATCATATC ATCGATCCGG TTACCGGGTT TCCGGCACAC ACTGATGTGG ATTCTGTGTC   11160

TATCTTTGCA CCCCGTTCCA CAGATGCAGA TGCGCTTGCT ACCGCCTGTT TTGTATTGGG   11220

GTATGAGAAA AGCTGTGCGC TCTTGCGTGA ATTTCCCGGT GTTGACGCGC TGTTTATTTT   11280

TCCTGAcaaG cgcGTGCGCG CAAGTGCaGG GATTGTCGAT CGCGTGCGTG TGCTCGATGC   11340

ACGTTTCGTG TTAGAGCGTT AGGACAGCAC GTGTGCTGTT CGTGTGTAAA AAAGTGTGGC   11400

GGACTGTCCT CATCATGGTG TGTGTGCAGG ATGCGTGCGC GGGGGTTCGG TCAGATGTCA   11460

GGGTGTAGGC AAAGATGAGC GCAGCGCTGA CAAGAGGTGT TGAGTGCACC CTTTACTCCT   11520

AGGTTCAGTG AGCTGCGTAA TTTTGAATCG AGGAGTACAG TGATGGAGAC GTTTTTTACC   11580

TCAGAGTCTG TGAGTGAGGG TCATCCTGAT AAGCTGTGCG ACCAGATTTC TGACGCTGTT   11640

CTTGATGCCT GTCTTTCGCA AGATCCTCAC AGTTGTGTTG CGTGCGAAAC TTTTGCCTCC   11700

ACGTCCCTTA TCCTGATTGG AGGTGAAATT AGCACGCGGG CGCATATTAA TCTTACCCAA   11760

ATTGCGCGTG ATGTTGCCGC TGACATTGGA TATGTAAGCG CTGATGTCGG TCTTGATGCA   11820

GCGTCCATGG CTGTTCTTGA TATGACTCAT CATCAGTCGC CTGATATTGC GCAGGGGGTG   11880

CACGGTGCAG GACTGAAGGA GTTTGCAGGA TCGCAGGGGG CAGGGGATCA GGGGATTATG   11940

TTTGGTTTTG CGTGCCGCGA·GACGCCGGAG TTTATGCCCG CCCCCCTCAT GTGCGCGCAC   12000

GCGgTTGTGC GCTATGCTGC CACGCTTCGT CATGAACGCC GTGTGCCGTG GCTGCGTCCT   12060
```

```
GATGCAAAAA GTCAGGTTAC CGTACAATAC GAGGGACATC GACCGGTACG TATCAGTGCG    12120

GTTGTGTTTT CTCAGCAGCA TGATCCGTCA CCTTCATACG AAACCATTAG AGAAACGCTC    12180

ATAGAGGAGA TAGTGCGTCC GGCGCTTGCA CCTACAGGTC TGTTAGATGA AAACACGCGT    12240

TTTTTTATCA ATCCAACCGG TCGTTTTGTC ATTGGCGGTC CCTTTGGGGA CACTGGTTTG    12300

ACCGGGAGAA AGATCATCGT AGACACGTAT GGGGGAATGG GCCGCCATGG AGGAGGCTCC    12360

TTTTCAGGTA AGGATGCATC TAAGGTAGAT CGTTCTGCAG CGTATATGGC GCGTTATATT    12420

GCAAAAAATA TTGTGGCAGC CGACCTTGCT GAGCGCTGTG AGGTGCAGCT TGCATACGCA    12480

ATCGGCGTAC CATATCCGGT TTCGCTGCGG ATAGAAACAT TTGGAACGGC GCGCGCATCT    12540

GAGTCACACA TCACACACGC GGTGAAAGAG ATTTTTGATT TAACCCCAGC GGGTATCGTG    12600

CGCACGTTGG ACCTGTGTGC GCCTCGGTAC CGCTCGACTG CAGTGTATGG TCACTTTGGG    12660

CGCGAACAGT TTCCTTGGGA ACGCACAGAc TGCGTGTGCG ACTTACAGCG TGCGGTGCGC    12720

CCGTTCGCGC TCTCTGGCCA GATAAAGAG TAGCTTCGTT TCTTTTTTGT CTGCGCGGGG    12780

CCTGTATCGT TACAGCCCTT CACTTTCTGC CCATGTTACG ATGATTGGCT CTAGGGAATG    12840

TATGGAAAAC CCAAGGGTAT GGACCTGCTG GTATTCATGA CTGTTGGGCC ACCGTTGGTA    12900

GGGGTCATCG TAGTGCGTGT GCAAAAAGTG ATAGATGGTG TCTTCTGCAT TGTTTTTGcG    12960

CGCGCGTAGG CgCAGACCAC GGCGTACTGT TGcACGGTTG AGCACCGTAC GAATGGCGGT    13020

CCGGACCTCT TTGAAGAACA ATTCGCGATA CACGCCGTAG TCCTTCCCGG TGATTCTTTT    13080

AACAATGTGT GCGATGGATT CGACCGTCTC TTCCGGTGGG GAAGATCGCT CTACGGTTGC    13140

AGCTCCCTCC TGGTGCACGG GAGAGGAAGT GACGATATCG GAATGTCTTG CCTCCGCGTG    13200

TGTTCCTGTG CGTGATGCAC GGGAACTGCA CGTGCGGACG TTCTTTGTGT TGTGACGGAA    13260

AACCCTTTCC ATTCCCGCTG TTTCATCTTT CTGTTTTGAT TCCCCACTAT CTTTTTTACT    13320

AATTTTCTCC GAAGGAAGGG CAGGTAACTC TTTCTTACGC GCACGAGTCC gTGCACGCGT    13380

GCCCGCCGTC TTGCGCGCAT GCGGAGTAGA AGATCGTCTG TGAGTCGCAG GCACTTTTTT    13440

ACCCTTTTGA TGATGaTgCT CCGCACGTTC TACCAGGCGC GCTTTTAACC AGTGCCAGTT    13500

TCGTTGCAAA TACAGATCAC CACTGATGCT GTAGTAATAC CCTGAGGCGA CGTGCGCGAG    13560

CACCAAGTGC AAAATATCTT CGTCTTTCCA TTGGTGACTC CGTCCAGTTT TAATTGAATG    13620

AAAGATATTG TAGTGATGTG TTTTATACTT TCTGCTGTAC AGTAGCACGG TGTTCTCAAA    13680

GGCTCGATCA TCGTATCTGT GCCAGTATTT GATTGCATAC GCAAGTACCC GATCTTCAAT    13740

TTTTGTAATA GTTTTCGCGA TTTCAATGGT GCGTACCTGT TCTGCGGTGA ACACCAGTTG    13800
```

TGAAAAATCT ATCGTTGAAG AGGTATCTGT ACCGTGAGCC TCTGTGCAAA AGCCGTAGtT   13860

GCGCGCGTGT GTGACTCCTG CCTAAATGTT CGCACAGAAG AAAGCGTGTC GGTAACGTAC   13920

ATGCGAATGA TGTCTGCAGC CTTTTCCATC TCTTTCTCTC GGTAGAGCAG CCATTGTGCA   13980

TAATGCGTGT GCTTTCTGGA GAAATGAACG GACTCAAAGC GGTTGAGAGA AATAGAACGC   14040

ACAATCTTTT CGCAAACTCG GATCAGCGTG CGCACGTCGG TCAGGTCGAG GGGCGCAACC   14100

GCGGCGGCGA AAAAATCAAT AATTTTCTTT ATTTCTTGGA TTTTCTTCTG AGCGAGCTCT   14160

GCTAGGGCAT CGGCAAGCAC TTCCTTAAAC CGCACGTCCC TTTTGCGATA ATAGGACATG   14220

AGGAGTACCC TACGCTCCTT CTGAGGGTAT TTCACCAAAA GCGTGCGGTT GTAGATATCC   14280

AGTGCCTGTC TTTGATCTCC AAGGGATCTT AGTTCAAAAT ACCGGTCGAT GTCGGCATCT   14340

TCGCTAAAAC TCAGCTGGGG AAACTCCGCC CTCAGTACTT TTTTATCAAG CTCTGTGAGT   14400

GTGCGCATGA CATCCCAGGT GCGTAGACAA GCTCTTTAGA AGAAAGGCTG GCGCGCAcGG   14460

GCCTTTtCGA AGCGGTGAGA AGAACTAACG CAAGAGGCTT AGAACGCTCT GcGTAgCCTG   14520

ATTCGCCTGT GCAAGCATAG CAGTCCCAGA CTGCACCAGA ATCTGGTTCT TGGTGTAGTC   14580

TACCATCTCT TTTGCCATAT CCACGTCGCG GATGCGAGAC TCAGCTGCCT GCAAGTTCTC   14640

TGCCGCGACA TTGATACCGG CAACCGTGTG GTCAAGTCTA TTCTGGTAGG CACCGAGATC   14700

AGCGCGCTGC TTATTAATTC TCTTTATTGC CTGATCAAGC GTACCGATTG CGCGGTTGGC   14760

CTTTTCAGGA GAATCGATAT TCATGACCGA CTCGTCACCT GCATCCCGAA TTCCCATGGC   14820

AACTGCAGTC ATGGTCCCGA TATACGCACG CGTGCGCTGG TCCATGTTTG CACCGATGTG   14880

GAACCACATG GATGCAGTTA CAGTGTTCTC CCCGCCTTGA CGCGCGAAGC GACCAGTGAG   14940

CATGTTCATG CCATTGAACT GAGCGTGGCT GGCAATGCGA TCCACCTCTG CTACCAACTG   15000

AGAGACCTCT ACCTGAATGT AGAGACGGTC TTCTGCGGAG TAGATACCGT TCGCCGCCTG   15060

CACACTCAGT TCGCGAATGC GCTGGATAAC GTCGGTGGTC TCCTGTAAAA ACGCCTCCGC   15120

AACCTGAATG AAGGAGATGC CGTTCTGCGC GTTTGTAGAC GCCTGGTTCA AACCACGGAT   15180

CTGGTCCGCA TCTTTTCAGA AACTGCAAGA CCCGAAGCGT CATCCCTGA CCGGTTGATG   15240

CGCAGTCCTG AAGACAACTT CTCAATGTTC TTCTGGACGG ACAAGTTAGT GTGTCCGAGC   15300

GTTCTTTGAG AGAACATAGC GCTCATGTTG TGATTGATGA TCATGAAGCA TTACTCCTTT   15360

TGGTGCTTTC AAGCGGACCA GCCGCCCTGG CATCCCTGCC gTTGCACCCC GTGCTTGGTA   15420

AGGGGTATCG GAATACGCCG GGTGCACTTG AGGAAAAAGC GGTGCGTATA TCTTGCGTAC   15480

GTGAGTGCTT GAACGTTGTG CAAATCGGAG GTAGAATCCC CGTCCTGTTG ACCTCTGCAG   15540

CAGAGTTACC CCGGTTAGGT TCGTGCGTGA GATAGGTTGC CGGTTGCGTC CGGCTGTGTG   15600

TGCACTGTGG ATTGAGTGGC TCTGTCCTTG TTTGAGCTTG TGCGCGGCGT AGCTGTACTT   15660

GGCGTGCACT GCCGTCTTAG CTTTCCACGG AGGGATGTGG GAGAAGATAA TTAGGGAATG   15720

TGGGGAAGGC GTATGAGGTG TATGAAGATA CCCAGGCAAC TGACGAGGCG TCGGCTACTT   15780

GAGAGGTTTT ACGCGCACCC GTGGGTGCTT GTTGCGGTGC TTAGCgcGCt GACGCTCTTT   15840

TTTGCAGTCC aGcTACGCAC GCTACGCTTG GACAATAATA ATTTTCGCTT TATCCCCAAG   15900

GAAAACTCGG TGCGTATCGC CGATCAGCGC ATCGATAGCA CATTCGGCTC CCAAGTTCCT   15960

GTGCTCATTG GTATTAAGCG TGAGTATACT TCCGTCGTTG ATCCTGTCTT TCTTGCGGAC   16020

GTGCGGTCGC TTATTGAACG CATCAGTGCG GTCCCCTTGG TGAGGGCGGA GAGTACTCTC   16080

TCACTCCTGT CTGCCGAATA CCTTGGTCTG CGTGCAGGAA ATATTATCAG TGAGCGTGTT   16140

GTTCCTGATG AGTTCTCCGG AAGTGCAGAA GAGGTACAGG GCGTTTATCG AAAACTTCGA   16200

GATTGGGATT TCTATGAATG TAGTCTAGTC TCGCGCGATC TACGCTCTAT GCAGATAGTC   16260

GTGTTTCTAG ACACCTCCAA CGAAGAAAGT AGTTCACCTG AAGCGATGGC AGCTTGTCGC   16320

GCGATCATAC GCATTCTCGG TGCGTGGAAA AGTCGTGACG CTCAGACTTT TGTCACAGGG   16380

GTGACTGTTT TTAACGAAAT GGGGAATGAG GCGTCGACGC ACGATTTAAC GCTCCTGGTG   16440

CCGCTTGTGG TGCTCATAAT AATCGTGGCG TTGTTTGTAT CGTTCGCCG CCTGGCGGkT   16500

ATCTTCTTGC CCCTTTTGAC AGTGGTCATA TCTACCGTGT GGGCCTAGG AGCTATGGCT   16560

TTGTGTGCCA TACCACTTTC TATCCTTTCT GCCATCTTGC CTGTAATTCT TATTGCCGTC   16620

GGGAGCGCAT ACGGCATTCA TATAGTTAGT GCGTATTTTT ACGGCGCCTC CTCGCGTATC   16680

TGCTCCACCC GGCAGGAGCA TCGCGCTCGC   16710

(2) INFORMATION FOR SEQ ID NO: 46:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1235 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

TCAGCCCGCG CACAGGAAAG TATAnGATCG GCACGTTTCC TTGCCTCCGC ATATATATCT   60

CTTTCTAACA CcTCTGTTGA ACGCAGCTCT TCCATGTAGT CTATACCTTT GCCCCGATAA   120

CCTGTCGAAT TGATTCACGC AGGGACGTGC GTACGTTCCC CTGTTCGTGC AAAGCGGGGA   180

CTTCTACGAC AAAGGGTAAC TCCCCCTGCA TCTGCCACGC ACGCACCTCC TCTCTCAGGT   240

AAGACATAGC GCGCTGGGTG AGTATGAGGA CCTGCGCGGG ACGGTCTCCA GAGGGGGGAC   300

GCGTGACCTG GGAGAATGCG CGCGCCGCCT CCTCTGGAGA ATGTACCACG AAACCGTGCA   360

CCCCTACTAA GGAAAAACCC AACACCAGTT CTTGCTCTCC AATTATGCAA TACGTCACTT   420

GATGACACCC CAGACGCACG TCAAAGCAAG ATGATAAGCA ATGCGACCAA AAAACCCCAC   480

AGACAAATGC CTTCTGCAAG ACCGATAAAG GGAAgTGCCT TTCCTGAAAT TTCAGGATCC   540

TcAyTCATTG CCCCCATCGC TGCAGCCCCG ATTTTACCTA CTGCAAGGCC TCCCCCAACG   600

CAAGCGAGCC CCACCGCGAG TcTGCGGCAA TGTATTTTAA GCCGCCATCT ACATGAGAGG   660

GCGGCTGmmT CTCCGCGtTa AGaAGACACG CACAAGCCAG CAAACnCACC CGTAAACCAT   720

GCTCTTTTCC AACCCATACT AATCCTCTTG ATATCCAAAC CTAAnCGGTG CGAAGACGCT   780

CCCACTTTTG GTAAAAAACT TTGAAAAAAA CTCGTAGTAT TGCAGCCGAA CCGCTTGAAt   840

GGCAACGATC AACCCTTCTA GAAAGATAAT GACTCCATTC CCAAACACGT ACACGAGTAT   900

GCCCCATAGT GAAGCGTAGC aCCAACGAAT TGCGTCATAG TAAACACCAC AAAACTTAGT   960

ACCGCATGGG ACAAGGCAAA GGCTCCCACG CGCAAAAAAC TCATGGAGTT GGAGAAAAAT   1020

CCCGACACCA CATCGACCAT TTCGATAACA CCGTGCATTA GATACATGCC AACACCTTCA   1080

GGAAACCACG GACGCACACG CTTGCACACA CGCTCCAAAA ACTCTTGACA AAAAtACCCA   1140

CGaGAGGCAC GCCCTTGCAA CCGCATCAAA GACCCCGAAT GGATTCCAAA AGTGGTATGC   1200

GCACTGCAAG GGCAACATGT ACCAAAAAAA GAGGG   1235

(2) INFORMATION FOR SEQ ID NO: 47:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 16636 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

ATTCTCnGCA CATGTnCCCT GACACTTCCG TAGCGGCTGC CGAGACCTGC AGCCAGAATA   60

ACTAGCGTGA CGTATTCGCT CATAGAAGTC TTCTAGCAAA ACGGAGCACG CCCCCGGgCC   120

ATCCCCGGGA AAGCGTAAGA GACGCACGGT GCACTTATGA GCGCGCACGC AAGGCCGCCA   180

TCGTAACGTA TTGTACGGCT TGTTAAAATG CGGCAGAAAG AACATGTCCA GCAGTTTTAA   240

CTTATCAATG GTCACCCGCT CCTGGATTGC AAGAGAGAAC AGGTGAATAC CCATCGACAT   300

454

```
GTCGTGCCGT  GACGCCATTT  GAGCGCCCAC  AATCACCCGC  GTTTTTTTAT  CAAACACAAT      360

CCGGATTTTC  ACCGGGTGAT  TGTCCACTTC  CATAAAGGCG  GGTAACTGTG  AGTCTTCAAA      420

ATCCGTTACC  TCCACCTCAA  GTCCCATGCG  CGCGGCGGCT  nCCTGCGTCA  CTCCTGTGGA      480

CACCATTTTT  AAGTCGTATA  TGTTGATACC  GTTGGAGCCC  TGCACCCCAA  TGCCTTCAAG      540

TGGGAATCCT  GCAGCGTTGT  GCGCCGCAAC  GATACCGCTG  CGCATCGCAT  TGGTTGCAAG      600

CGCAATGTAA  GAAGTTTGTC  CGAGCGAATT  GTCAAACACC  GTTGCACAAT  CGCCAATTGC      660

GTACACGTCT  TTCACACTTG  TTTCTTGTTT  TAAATCTACC  gCATACGCGC  CATTGGCAAA      720

GGTTCGCACC  TCGTTCTTTC  CCAGTGCAGT  ATTGGGGCTA  AAGCCAATGC  ACACCATAAC      780

CATGTCTGCG  GGGTACTCAC  CCTTATCTGT  CACCACTGsC  ACTACCTTTC  CATTGCTGCC      840

ACGGAGTTTT  TGCACTTTCT  GGCCAAACGC  AAGGGTGATG  TGATGGTGCG  CAAGTTCTCA      900

TCCATGAGTG  CACGGAAGGA  TGCGTCGTAA  TAATTGGAAA  GACTAGAGTC  CATCGCATCG      960

ATGAGCGTTA  CCTTTTTTTG  GTGGCGCTCA  AACGCCTCTG  CAAGCTCCAC  GCCAATGTAC     1020

CCGGCACCGA  TAACGGCAAT  ATTCTTAATG  GAAGGCTGCT  CGAGTTTTTT  AATCACCGCT     1080

TCAGCATCCT  GAAATAGCTT  AATGCGCTGA  ACATTCTCCA  AATCCATGCC  GTCGATTTTA     1140

GGAATGATAG  GCAGAGAACC  GGTTGCAATA  ATAAGCTTGT  CGTAGGACTC  TGCGATTGCA     1200

GAACCGTCTC  GTGCAGTCCC  GTACACCTTT  TTTGAGGCAA  AATCGATACG  GGTGATATCG     1260

CTTTCCATGG  AAACGCGTGC  ACCCTTTTTT  TCCAATTGTT  CTTTATTTGC  GTAGAATAGA     1320

CCCTCCGATC  CACGGATTTG  TCCGCCAATC  CAAAGAGCCA  TGCCGCAACC  AAGGAAGCTA     1380

ATATTATTGT  TACGGTCAAA  GACTACCACT  TCATTCGTGG  TGGTAAGGTC  TGTGAGGCAA     1440

TTGACACAGG  CGGTTCCTGC  GTGGTTCGCT  CCGATAATGA  CGATCTTCAC  GGCGTCCTCC     1500

TTACGTTTGG  CATTGTAGTT  CAGGGAAAAG  ATTTTTGTAC  AGGCGCCTGA  AAACAGCCGC     1560

GGTTGTTTTG  TTCCCAAACG  CGATAACTGG  GAGAATGTTA  TTcTGCGGTG  CAGGTGGTTT     1620

TTCTTCAAAA  ATGAGCGCGT  GGAGCGCGCG  AACCCGGTCT  GAGTCTTCTT  TTTGCACGAG     1680

GATATTCCAC  TGCGTTGCCG  CTTGCTGTCT  ATCCTGCATG  GTTAGATACA  GATGAATCAA     1740

CTCAATGCGC  ACCTCAGTAC  ATTGGGGCCA  CAGGGTGCAC  GCCTCTTGGA  GCGTAAGCTC     1800

TGcACGGGGA  AAATCACCAA  GTTGAGCAGC  AACGAGTGCA  ATCAGCGTAT  ACACATGCGC     1860

AGTAAGTTGT  ATATCTAACG  CCTTCACCAG  TAAGGCGTCC  GCTTCATGGA  GTAAATGCAG     1920

TTTGATGCTG  TTTTCTGCTT  TGTGCACGAG  CATCTGCACG  TTCTGCGGcT  CCTCCCGGAG     1980

CGCTTTGTCG  TACCACGGCG  CTGCGTGTTC  ATAGCATCCG  TCGGCGTAAA  AGGCGTTGGC     2040
```

```
AAGAAGATGA AAAGCCTGAC CACGTTGCGC AATCACCGCC gCGTCTGCGC GCTGCGTTTC    2100

CACTTCAAAC AGGGGAACAC AGCaCGCTAG CGCACCGCgc TGGnCGCTGT AGTTCTATAT    2160

AATTTTCCAA CACTACGAGT TCGTGCGGTA ACAGCGCACG CGCGCGTTGT AAAGACTGCG    2220

CTGCAGCGTc GAAGTTTTGC TCGCGCAGAG CCTTTTGCGC ACACAGGTTA TGCAGCCAGC    2280

CGTTATCCGG ATCGAGTGCC AGGGCGCGCG CAAGAGGCTC GTCACAATCT TTTTCGCTTA    2340

AAAAAAGAGA TTCAGCGTAT TTAAAGTGGT AGAGCGCACA GTCCGGCGCA AGTGCACAGG    2400

CTCTTTGAAA TGCATCACGT GCGCGCTGTT CGCACGCTGC TGCAGCAGGA GCGTCGTGTT   ·2460

CGTGCGTGCC CTGCGCTTCT CTCAAAAGCA GGCCATATAA GTACCAGACG GTAGCATCAG    2520

CGCTCCtGCG CGGCAGAGTG CGTCAAAcTG CGTGTGTGCC TGACGGTGCC GTCCTGTTGC    2580

ATAATACAAT TTGCCTGCGA TACTGCGAAC TTCTGTACGC TCAGGATCAA GACGGCGAAG    2640

TGCGAGTACC ACGCGTTCTA AATCAGCGTA GGATTCCTGT GCTAAAAAAA GCCGCGCCGC    2700

ATGTAAATAC GCCTGAATTG CTTGCTCCTT TTCCCCCAGT AGGGAAAGCT CCTGCGCCGC    2760

ATGCAGCGCA AAGAGTCCCT GGTGAGGGTC CAAACGAAAT GCACGCTGAT ACGCAGCAGC    2820

AGCGTCCTCG TGTCTATTTT GTGCAAGTGC AAGGTGACCG CATAGATTGG AAAGAAATGC    2880

ATCGCGCTCG GCTACGACAC GGTGTGTGGC AAGGAGGTGC GCGAGTTTCT CGTATGCGTT    2940

TTGTTGGTAC AGCGCACCTC CTAAGGCGAG AAGTGCCCGC ATACAGTGAG GGTCTTGCAC    3000

AAGCGCTGCG TTAAACGCTT CTTCTGCCTC ATCATAGAGG TGCTGGGCAC ACGCGATGTC    3060

TCCTGCAAGA AGATACTCGC GTAAACCAAA GGTACATTTG GCTTTTAGTT TTTTCATCTC    3120

TGcACGAGCG CaCrCTGCGc GTcCCATAGC GTACAGGGAT TCTGCGAGAC GCAGgTGCGC    3180

GCTTGCAGGG ATACGCGCAG gTTCGTTGGC GCGCAAGCAT GCAGTTTCAA CCATGAAAGG    3240

TTGCGATTCC AGACTGAATG CTGAAAGTTC AAAAGAACGA TGCGCATGGA CACCCCAGTT    3300

CTGGGCAGCA AAACAAACTT TCCCTGCCGC TTGTATCGTG TCATCTTCAA TTTGCGCAAT    3360

CCACTGATCA TCTACACAAA GAGTAAAGCT TGTGCCAACT GCAATGAGGC TCAACACGCG    3420

CACTTCATCG ctGGCGCCAG TGTCAGTCCA TCCGAGCAGG GGGAGGGGGG TGTTGTTGAC    3480

TACCGCGTCC AGACGCAGCC ATCCACCGTC TGAAACAAGA AGCGCGTAAA AGGTGCTCTC    3540

ATTCAGATAA CGAAACAAGA GCCcTGCGGC GCAGGTACCT GCTCGCTCGG GCACCGCCTC    3600

GCCCATTGCG CTCGGGTCGA CGGCAACTGC CTCCGGAAGA CAGGCAGGCC GAGAGGTCGA    3660

ATCAAGAGAA GCAGGTACTT CTGGCTGTTC TGAATTAGGC AGCGCACGCG CCTCGCCTGG    3720

AGGGTGAGTA CCCGGGAGAA AGCGGATGCG CGCAGTAAGC ACGCAGTCCT TATAACGAAA    3780
```

```
GACGGGGTTA GCACTCCACG CATAGAGGAA CTTACGTCTG AGGTGGAGCG TTAACCCATG   3840

CGCGCCACGT GCAGTCTCAT AGCCGTCCCC TGCCTCCGCG TGCCAGCGTG CATGTTCCGT   3900

AGAAGAAAAG TCAGCGCGCC AAAACTCAGA CACAATTTCT TCATAATCTA CAGGAGGTGT   3960

TACACGCCTT TTCAAGAACT TTGCTTTGAA ATATTGAAAA TACAGGCGCG CACGTCCCAC   4020

GCCCAAACTC TACGCTAATT TTCCCCAAAG TAAAAGGGGA AGGGTGAGTC ACAAGACAGC   4080

ACACAGGTGA TCACAGGGAG TGCGCGTCTC CGGTTAGGGA AATAAGAAAT GTGGTATGCT   4140

CCGCCTGTAC GTTTGGACTA TGGTGCAGAG GATAGGAAAA GATGCAACTG TACACTCCTG   4200

CGCCTATCTG GTACCCTCTG GcATAGTTTT TACCTAAGGA GCATTTCAAT GGCATTACTT   4260

GACATAAGTA GCGGGAACGT CCGCAAGACT ATCGAGACCA ACCCTCTGGT CATTGTGGAC   4320

TTCTGGGCTC CCTGGTGCGG TTCGTGCAAA ATGCTCGGTC CTGTTCTGGA GGAGGTAGAA   4380

AGCGAAGTCG GCAGCGGTGT TGTTATTGGA AAACTGAATG TCGATGACGA CCAAGATCTC   4440

GCCGTTGAGT TCAATGTGGC GAGCATCCCC ACGCTTATTG TTTTTAAAGA CGGGAAAGAA   4500

GTCGATCGTT CCATAGGCTT CGTTGATAAG TCAAAAATTC TCACGCTCAT CCAGAAGAAC   4560

GCCTAAGGAT ATTTCTTTCG TACGGAGTGT GCTACCAGCT CATCAGCAAA GCGATGCAGC   4620

CGGTGCCTAG GGGAACAGTT ATATTGTCGT AGTCTTTGCA GGGGAGCAGT TCGATGAGCG   4680

CCGCGCAGAC TCCGAGCCCC CAGCtGCGCG CGCCGGAGnT CCCAGCGCGT GCACACTCAG   4740

TGCCGTTGCT ATGCAGCACA CGGCACTGCC TACGGGTGTT TTCCCCTGCA CGGAAAAGAG   4800

CGTCGTGGTC CTTCCCCTTC TGCGTCCGGG GGAATCAGCG GAATTTTTGA AGCGCAAGAA   4860

AGTGTATACA GTACCGGCGA GGCTTGCGCA TCCGTCTCCG AAAGCGAGGG CGCAAATGGC   4920

GGCTGCTGCG ATAGGTGGGG GAAAGAGAAG GATCGCGCTT GAAACACCGA TTGCAAGGGT   4980

GAGTGGCCCT CGCACAAGAG AACGTTCGTG TTTGTCCCGG TCTCGGGCCG CCGCTCGGGT   5040

GAGGAGTGAG ACAAGTGGCA ATGTTTTTCC ACGCAgTCTC CACCGCTCGG CAACATAATA   5100

GCCAACGCCG AGCGTGCCAA TGGCGCCGAG CGTAAGGGGT TTGCTCCATG CCGCAAGCAC   5160

AATGCTCAGT GCTGACGACA GGTGTATGCC CTTTCTGAGG CATTCGCTTG CAAGGCGTGC   5220

GTGCATGCTT TCTACGTGAG CGGCGCGGCG AGTACGTGCG GCTGGGTTAG CGCCGGCCTG   5280

TACGGCCCCG ATCCTGGCTG GATGCTCGCC CCGCCGGCAG GGGGCGGCTG TCTGCAAGnT   5340

GTTGTTCTGG TCCTCAGCGT CCACCGGCTT CCGTCTTTTC ACAGGGGCCA GCGCAgTACT   5400

ACGCGCGCTA CATTGAAAAA TACGAAGAGA CTGGTAGTAT CTACGATTGT GGTGATCAGC   5460

GGTCCTGCCA TGATGGCGGG GTCTAAGCAG AGCTTTTTAG CAAGAATAGG AAGGAGCCCT   5520
```

```
CCGGTCAGCT TTGCCGCTAC TACGGTAATC ATCAGTGTTA TTCCCACGCT CCCACATAGG    5580

GCAAGCGGCT TGCCATTTAT GAAGTACGTT TTCGCCAGGC TTGTTGTGCC TAGAATGCCG    5640

CCAACTGCCA GCGCAACTCG CAGCTCTTTG TACAGCACCC GGAGCCAATC GTGTAGCTGA    5700

ATCTCACCCG TTGCAAGGCC TCGGATAATA AGCGTAGAAG ACTGACTGCC TGAATTCCCC    5760

CCGGTGTCCA TGAGCATAGG GATAAACGTG GTAAGCGCGG TGGCAGTCAC CAACAAATCC    5820

TCGTATCCGG CGATGAGATT TCCCGTGAAC GTTTCAGACA CCATAAGAAG TAGTAGCCAG    5880

CCGATACGGT GTTTCACTAG GGTGAAGACC TCAGTTTCCA GGTATGCTTC ATCTGAAGGC    5940

TGCATGGCCG CCATGATCTG GAAATCCTCG GTAACTTCCT GCTGCATCAC GTCCATGATG    6000

TCATCGACGG TGATAATGCC AATGAGCCGT CTTTCAGTAT CCACCACAGG GAGTGCAAGG    6060

AAGCCATATT TTTTAAACAC CGCCGCGACT GCTTCTTGAT CATCATGGGT GTGTACAAAG    6120

ATGCAGTCAC GTTCACACAG ATTCTCAATC AACAGATCTC CCTGACTAAG CACGAGCTTT    6180

TTTAAAGAAA TGACACCGTG CAAAAACCGG TTTTGGTCAA TGACATAGCA CGTGTACACG    6240

GTTTCTTTTT TTAATCCGGT TTCCCTAATG CAGCAGAGGG CATCGTGCAC GGTCATCTGC    6300

TTCTCTAAGT CTACATATTC AGTTGTCATG AGGCTTCCTG CAGAATCCTC TGGATATTTT    6360

AAAAGTTGAT TGATAACTTG ACGTTCTGTT TCACCCGTTT GTGCGAGAAT GCGTTTCACA    6420

GCATTTGCAG GCATTTCTTC TACCAAATCG ACTATATCAT CCATGGCGAG TTCTGCAAGA    6480

ATAGGTGCAA GCTCTCTGTC TGTAATGGTG GCGAGAAAAG CAGATTGTTT GCTGCTTGAA    6540

AGCTGCGCAA ACACATTTGC AGCGAGGTTC TTGGGCAGCA TTCTAAATAG CAACAACGCC    6600

TGTGCAGGTG ATTGCATGTC CAAGACATGC GCGACGTCCA CCTCGTTCAT CTCGTTTAGA    6660

TTTGCAATAA GTGGTACGTA ACGCTTGTGC TCGAGAAGCG TCTGGATTTT TTCAAAGTTC    6720

TCGTTCATAG CCAATTCCCA CGCGGAGTTc CGGAGTATAC GTGAGTTCAC CTCTGTTCTT    6780

CCATAGGTGC ACGTCCCGCA CGAAAGTGAC TGCTCCTGCT CCCAAAGGCC TGGCGCCACC    6840

TAGGCAAAAC GCACAAAGC AGAGAGGGAG CGAGAGACGC GTTCGTTCAG GCAGATCAAT    6900

CGATGGAGTT CAAAATTCGA GTGAGCCTGT TTTTATCGTA TAGATCCATA GGTCTGCCCT    6960

CTGCAAAACG GCGCGCCTCA TCAGTAAAAG CACCGGCGGT CATGCAAATG CCCTTGCCAG    7020

CTTTTAGCTC TCTGATACGT CCATGCAGAT CGCGCAAAAC GAGTTCCCCT ACTGACCCCT    7080

GAGAGCGAAA AAAACGGAAC AATACGAGGT CGGCCCACTT TGGCGTGTCA ATTTCAGAAA    7140

CTATATCTGT GTGAGTATTC AGAACCGATA TGTCCAAAAT CTTTACCCGC GCGTGCGGAA    7200

AGTACTTTGA TACCACCTTT CTGCATAATC CAATGAACTC ACTCTGAGCC GCTATCATGT    7260
```

```
ACAGGTGTAA GTTTCTATTC TGATTTAATT CCTGATAGTA CGTAATGAGC GCCGTGACAT    7320

CCCGGTACTC AGGGTTAAGC TGGTGAATCT CTCTAAGCAT AACCAGGGCT CTTCCCAGGT    7380

TCTGGGTTTT TATCAGTGTC TGCGCGAGAC GATAACGCAG CTCGTTTGCA ACGTCTGAGG    7440

GTATATTGTC ATGCTTCAAC CCAATTTCAA AATCCTCTGC AGCGTCCTCC AGCTGGTTCA    7500

GCTTTGTCTT AATGGTCCCT GCATACAACG CAGCCCGCGG CCCAACGAGA GGATCAACCC    7560

TCAGATGATT GAAGATTTTC AATGCACGGT CGTGTGCATT CGCCTCATAG AAGCACTCTC    7620

CCATTACGAA CAATGTTTCT TTGTCTTCAG GTTGGAGGTC CAGCGCCTTC TTCAGATAGG    7680

GAAGGGCTTC CTGATACTTT TGCAGTTTTT GAAACGTGTA ACCCGAgcaG sTGCGCTGCA    7740

GGTGCACTAT CCGGTTGCGA TGTTAGCGCT CTTCTCAAAA GAGGAACGAC CTGCTCATAC    7800

TGCCGGTCGA GGTAATAGAC CCTTCCCAGG TTGTAGTTTA CTTCAAAATG GTGGGAGTCG    7860

ATGCCTTGCG CAAGGAGCAG TCCCTTCTTT GCCTCCGCGT ACTTTTCTGT TTTTACGCCG    7920

CAAATCCCGT ACCGGAGTAG GATCTCAAAC TGCTCCATCT GTGAGCTCGA GCTGcTGCGC    7980

TCCACGAGTG TCGCGTACAT CGCGAAGGCT TTTTCCCACT CCTGATCCTG GTAGTGTATG    8040

TCCCCCAGTA CCGAAAGACC CCGAATGTCA TACGGGTCCT GCGCAAGCCG TCTAGATGCT    8100

TCCCTCATGA GAGACTCTTT GTCTTTTCTT CTGCTTGAAC GCGTCCCCAC CTTGCTCGCT    8160

ACCGTGGCAA CAAGCATGAT ACTCGAGAGC GCAAGCATAA CGACAAAGAA GACGATAACG    8220

GCAGAACTCA CTCGGCGCAG TGTGCTAACA AATCATTCAG CTGTCAATAG ACTCTGCAAT    8280

GTGCCGGTAG gCTGGaAACA GAACTTAAGA AACCTGACTG CATATCTTTG CAACACCTGC    8340

GCGCCACTCC AGCAGTGTAT TTTTTGTTTC CCACTCAATG AGCCGAACAA TACTAAAGCA    8400

GGTGGTACTC GCTGGGTTGG GAAAAACAAT AACTGCAAAT TTCCCTTGGC CAACATACAC    8460

GACATGTTCG TGCGCAGCCA AACACTCAGA CTCCCTTAAA AAACTGAGCA CACAGTCAAA    8520

ATGAACTGGT CGACCGGAAT CCTTGTCACC AAGAGCTACG GGAAGATCTT TGATAACCTC    8580

TTTGCAACGC GGACACACAT ACTCGCAAAA CTCCAGCTTT GGCGAAAAAA CATCACGCCT    8640

CGGCGTGCGC TTCCTTTTTC TTCTACGGCC GCTTGACTGA ACGTCAGTCA CACTCTATTC    8700

CTCTCATTGA CAATAACCGC TTACTGAGAA TGAGCGCGAG CCTTTGAATA GCCCCCTCAA    8760

TATAGGACTG ATCTTGAATC TTCTTTTCCA GTTCGTGAAT TGTTTCAGAG CAAAACACcT    8820

GCGGGAATTC GTCTTCCTCG AGAAGGAAAA ACGGAACCTT TCATACCAAA TCCTCAAAGG    8880

CATGTGGCAA GTGCACGATG TCCACATCCT GCCTTCGAAA AGGATCGGAC CGCAGCACGA    8940

TGACGTCAAA GCGCGCACAC ATGTGGTTGT ATTCTCGAGC GCTGGCAAGA AAATGTTTAG    9000
```

```
CGGTTTCACA GATCCTTTTC TGCTTGCGCT TTCCAACAAT AATTGCTAAG TCGGCGTAAC    9060

TGGTACACCG TAGCGTCTTC ACTTCTACGA ATACTATTGT GTCATCCtGc TGCGCAATAA    9120

TATCAATTTC ACCTGTTGCT CTGCGCCAGT TTCGTGTGAT GATAATATAT CCGCGCGTCG    9180

CTAGCCAGCG CGCCGCATAC GCCTCTCCAA ATGCACCGAG TAACTTATTG TGCTTTGGCA    9240

TAACTCCACG GACTCACTTC TCCTACAATG TAATTCGTTT TATTGACCGC TAAATCAATG    9300

AACTCACCGC TGTCGTCAGG ATGGGTTACG TCGTGCTCGT CGATAACTAT TCTTAAACAC    9360

GGACGCAACG GCGCCGTGGC CGTGGTCTTT ACCACGCGCG CAACTGCTGA GTTATTCAGG    9420

AGCACCAACG ACCCAATGGG GTAAATACCG ACGCTTTGAA TCATTGCCTT GATTACGTCT    9480

GGATCAAAGC GACGTGCGTT GTCAGCAAGT AAGCTTTTCA TTGCTTGATA GCCACTGAAC    9540

GGTTTGCGAT ACGACTTTGG CGCAAGCATT GCAGCAAAGT TATCAGTAAC TGCAAGAATC    9600

CTCGCACCGA TGGTAATTTT GTTTCCAGAA AGAGACTGGG GATACCCTTT TCCATTCCAG    9660

TGTTCATGGT GCTGGAGTAC ACTCAGTCCA ACCGAGTTCG GGTATTTGAG CGTGTTTACG    9720

ATGTAGGAAT GTGCGTAAAT GGTGTGCGCG TCAACTGCCT GCTGTTCTTG AAAATGCAAT    9780

CTCCCCGACT TCTTCAAGAT GTCGGCAGGA ACATGCTGCA TACCGATATC GTGCAGGAGT    9840

GAGGCAACGA CCAAATCAAA TATATCTTTT TCAGAAAAAC CCAAATGCTG CGCAACGATA    9900

ATGGAAAAGA TAGCCGTATC TACTGCAGGT TTTGCAAATC GAAATCCTTC GATTTTGTAT    9960

GACAGCACTA AACTGACAAA CCCGAGTGTA TTTGCTCGAA CTAATTCTGA AAGGCGCTTT    10020

GCAAGCATGT CCGCAGGCGC GCAGGAAGTG TCGTGTGCGC ATTCATCTTG GCAAAAAGCA    10080

TGTTCAATTC CTGAATAAAG CCAACGTATT CTTCATGATA GTGGGGATTG ATACACACCT    10140

TAGGAAGGAG TTCACAAATA TCCTTCAGGA TATCTTGAGT ACGGAgTTCT CCGGTGGAAA    10200

CGATATCGTC TTCAGGATCA AGCAGTTCTT CTGCTGCAAg cTCTTCAAAT TCTGCTAATT    10260

CCTCCACGGT AGAAGAGGGA ACTGACTCCC CTTCAGCCAG CACCCTACCT GCGGTCACAA    10320

CGTAGGGAAT ATTCCAATCC TGGAGCACCG TCAGCTCTCG AGTGCTTACC GGCTCCCCTT    10380

CCCTGAGGAG GAGGTTTTCT CCGTCGTCGA AGAACACGGG CTCGGAAAAA CACATTCCTT    10440

CTTGCAGTTC AGATACATCG ATTTTTTGTG ACATGGACCT GTACCTCTTT ACCCGCCTTA    10500

TCTTCGGCAT CGGTGCGCAC GGGTTAATAC CACGTCTTAT GCACACACAG CGGTAACGTT    10560

ACTGTTGTGT GTACAAAAAG GCAAAGATTG CAGAGACAAC CTGATAGGTT TCCGGTGGGA    10620

TACACGCaCC TATCCTGTGC TCAGACAGAA CACGCGCCaG AAGTTCGTCT TGCACCAAGG    10680

CGATATCAAA CTTTTTTGCA ATTTCAACAA TTTTTTCTGC AATGGCGCCC GTGCCCGAAG    10740
```

CGACAATAAT GGGCGCCTTA TCTCCCGTTG CATAGGAGAG CGCAACCGAG CACGCACGCT   10800

TTCTTTTCAT GGTGGCGGTA GTGTGCGTCT GCTATACAGA AACGTCAATC CCTTTGAAGG   10860

GGACCGCATC GTTTGCCGAG TCTGCATGCG CACCATACCG TACGGATAAA AAATCAATGC   10920

CsCGTTCGCG AAGGAGGGCA CACAGGCGTG CAACTGTTTT TTTCtGscTG TGCGCGACAG   10980

TTCCCGTGCC GCGTGTCTTT GCACCGTGAG TACGTTATTC TGCACGCAGA AGACCCATTC   11040

TCCTGCAGTG TTACACGCAC GTACGCGCAg cTGGgTGCAT GTTTTTTTGT GGAGGTGTAA   11100

TAACAGCGAG AGAGTGCCGC GCCACACGGT GTGCGCGCCG CgCCGCTCAA AGGGTACAAG   11160

AAGCCAGTGG AGTGCTGAGT GATATGTGTG GTTAACGAGC GCAAAGAGAT CTCGCTCCGT   11220

GTCAGAGTCT GTGTATGCGC CTGATTCCCC AAGAAAGGTG CTCAGCATAC GGCGGAGCAT   11280

CGCTGCATCA ACGGGAATGT TTCGGTCGCC TAAGATACTC GCAAGGAACG CGGCGTGTGC   11340

TCTTTTTTGC TCAGGAAACT TTTTTAGCAA AAGCGCAAAG CGTTCAATGA GCTGCGGCTG   11400

CAGCGGCACA CACAGGGATG TATGCGCGTG GATAAGCGCT GCAGCTTCAG GAGAGAAAGA   11460

AACACCCCAG CGCTGCAAAA AATGCGCCGA CATATCCTCA GCAGATGGAG GAGTGCTCGT   11520

GCACTGCGGG TGCAGGAACA CCGTACCCGC GCGGATAGAC GCGCGCAGAA ACAGGACTGC   11580

GCCTTTGTGC ATAGGCTGCG GTACACGCGC GCGCACCCGT TCACCGTTAA TGGCGACAAG   11640

CGCACGGCCG GCGTGCGTGC TGCTGAGAAT GCGGACGCAC ACGAGCGCCC CTTCAGTAAG   11700

GGAAACCGTG CGCGGCACTT CGGTGAGTAC TACCCGAACA GCTCCGTTCA CGCGCTGGGC   11760

GACGGCTTTT TAACAATACG TGCTTTGATA CGGGCGGCCT TTCCTATTTT TTCTCGGATG   11820

TAATAGAGcT TCGCGCGTCG CACCTTTCCT GCACGTAcTA CGTCGACCCG CTCGATACGG   11880

GGGGAGTGGA GCGGGAATAT ACGCTCCACT CCAACGCCAT AGGAATTTTT GCGCACCGTA   11940

AAGGTGCGCC TGACGCCGCT ATTTTTAAAA CACAGAACGA GCCCTTCGTA AGCTTGGATG   12000

CGCTCTGTTT TTCCCTCCAC TATTTTGAAA TGCACACGTA CGGTGTCCCC GACGCGGAAC   12060

GTTTCAGCTG GTTCCTTTCG CTGCTGGTTT TCAATTTGTT GGATGAGGTG GCAACTCATA   12120

GTCTAACTCC TTAAGAAGGG ACTCAGCCTC TTGAGTCCAG GCTGCAGACG CACGCGCAGC   12180

gctGaGGAGG TCAGGTCTAT TCCTTCGTGT TTTTTCGATC TGGCGCGCAA gcCGCCACGT   12240

GCGGATATGC GCGTGGTGAC CGGAGAGAAG TACAGGGGGA ACGTCCCGGT TGTGAAAACA   12300

GCGCGGCCTG GTGTACTGCG GGTACTCCAC GAGACTGTTT ACAAAACTTT CCTCCTCGAG   12360

AGATTCATGG CGGATGACAC CGCTAACACA GCGGCTCACC GCATCGATGA GCACGAGCGC   12420

GGCGATCTCT CCTGAGGAGA GAACGTAGTT CCCGATGCAA AtTCGTCGTC GACATACTCG   12480

● .

```
TCTATGATAC  GTTGGTCAAT  TCCCTCGTAT  CTGCCGCAGA  TGAGCACGAG  GGCACGTTCC    12540

TGTGCGAGTG  AGCGCGCATA  GCCTTGCTCA  AAGAGCTTTC  CAGAGGGAGT  GACGTACACG    12600

ACGCGCTTTT  TGGGAGCGTC  TACTGAATCC  AGGGCCTTCC  cTAACGGTTC  TGAGCGCATG    12660

AGCATGCCAG  GTCCGCCCCC  GTAGGcGGGG  cGTCACAGTG  TTTGTGTTTG  TCGTGCGCAA    12720

AGTCACGGAT  GTTGACAATA  TTGTAGTGAA  TGATCCCGTC  GCTCACGGCG  CGCGCCATTA    12780

TTGAGGTGGA  GAAATAAACC  CGCGGGATGG  CGGGAAAGAG  AGTCAGTACG  TCAATGTTCA    12840

TTCGAGAATC  CACCGCTGCA  ACAACTCGAT  TTTTTTTCGA  CCAACGTCCA  CGTCCCCAAT    12900

GAAGGTCCGG  TGAAAAGGCA  CATAGCAAAC  GCCACCATGT  GTCCTTTGAA  CCTCTAAAAG    12960

GGAGCTACCG  CCCCCTTCGA  CAACGCTCAA  GACAACACCC  ACCGCCGAGC  CCTCGAAAAC    13020

GAGTTCACAA  CGACACAGAT  CGGCCAGGTA  AAACTCCCCA  GCGCTAAGCG  GACAAGCCTC    13080

GGCaCgCGGT  ACCCGCAGCT  CTGCTCCTAC  AAACGTCCGA  GCGCACTCTA  CCGTATCTAC    13140

GCGGTGGAGC  TTGAGCAGCG  CGTcCTGCGC  ACGTAGGAGA  ACGTGCTCTA  CCATGTGGAC    13200

GGCCTCACGC  GGGAGAGCAC  AAGCGAGGGT  GCCTGAGGAT  CTGCTCCGTG  GAGGAGCAAG    13260

ACAAACCTGC  TTTAGTGTGG  CAAGATGTGC  ATACTCACCC  GAGAAGCTCT  TGAGCCTGAG    13320

TAAACCCGCA  ACCCCAAAGG  TGCTCACGAT  GCGTGCAGTC  ACAATTCTAT  CCATAACCCA    13380

CCACACACCG  CCTGCAACAA  GGCCGCAAGA  CGGAAAAGGA  GCCGCTAGTC  GATGATCTCT    13440

AAAGCGTAAC  GCGTCTGAGA  AGCGTGCGCA  GACGCAGAAA  GGAGCGTkcG  CAGAGCGCGC    13500

GCAATTCTGC  CGTGCTTGCC  AATGACCTTC  CCTACATCTT  CAGAGGCAAC  ACGTAACTGA    13560

AGGATCTCCA  ATCCCTCCCC  TGGAGACTTG  GTGACGGTAA  CCTCCCCAGG  ACGATCCACA    13620

AGCGCCCGCG  CAATATAGGC  GATTAGCTCT  TCTTCCATCG  TGACCATCCC  CTTGCCTAGA    13680

CGCCCTGCCC  CCCTGGGGAA  GAAGGGATTG  GAGCGGCGCA  GGAAACGGAC  TCTACGTGcG    13740

CCAGATCGGC  AGCCTGCTGT  GAAGAAGCAA  CACGGCGCTC  ATCTGAGGCA  AtGCGTTTAG    13800

GACAGAACCA  CGCCTGGACT  GCAAGAGCtG  CGGACCGTAT  CCGAGGGCTG  gCGCCGCGyT    13860

CAAGCCAGAA  GCGCGCACGG  TCAAGGCGGA  AAGACACCTC  GGTACCCTTT  GGGGCTATGG    13920

GCTGGTAAAT  ACCCAGTTCT  TCGATTGCCC  TGCCATCTCT  CGGCTCGCGC  GCGTCCTGAA    13980

CTACGATTCG  GTAGTACGGA  CGCTTCTTAC  TCCCCAATTT  TTTCAGTCGG  ATCCGTAAAC    14040

TCACTCTGTC  TCCTCCGCCT  GAGACACACG  CaGGTGCGCT  CATCCTTTCT  AAAATTATCT    14100

GTGCTGTCAA  GTGTCTGAGC  ACGTAACGGG  ACATGGAGAA  TAGATTACAG  AGGAGCGGCA    14160

CGTGACGCGT  CATTTCTGCG  CTGTAACGGT  TGTATTGGGG  GAGAAGGAAA  ACtGCAGTGC    14220
```

```
AGCGGGTGTG CCTTTGTCTT GGGTGTTCTT ATTCAAAAAT GATACGCACC TGGGATTTAA    14280

AGGTAGTACC TCGCTCAAAC TCGTGGGCGA AAAGTTCAAA ACTCGCTGCC CCAGGAGAGA    14340

ACACGATCAC GCCTGGGCTT TTCTGCCGCG CCCGGAGATC CTGCAGGAGT ACCTCAAGGG    14400

AAGTAAACGG TCCGTAAAAA GGTACCTGTG CTGCATGAAG AAGTGGTTGC AACCGTGCCG    14460

TAGCACTTyC TGCAAGCAAG TACAATGCGT GCGCCTTGGC TGCTGCCTGT GCCAAGGGTT    14520

GGTAGTCTGC ATTCTTATCA GTGCCGCCAA CGATAAGAAC CACGCTTTCA TCAAAGGCTT    14580

CCAGCGCTGC AATTGTTGCC TCAGGTACAG TGGATGCAGA ATCGTTATAA AAACGCAGTC    14640

CCCCCTTTTC GTAAAAAAAC TCTAGTCGGT GTTCGATGCC CGTGTAGGAC TCCAGCGCTT    14700

GTGCAAGACG CCGGGTGTGC TCTTGGAAGG GACTGTGAGC GGACGGACAA GCGTAGTCTG    14760

GGGGGGACGC GTGATTCnCG TACGCGGGGG AATGGGAGTG TGCGCAAAAA CACGGGGGGC    14820

AGGAGGAAGG AGGTAGAGAA TGCCCTGCGC GAACAGGAcG CTGCAAGCGC TGCACTCGCC    14880

ACTTGCGTTT GGAGGACACG GCCCGGTACG TGCAGCTGCG GTGGGATGAG CATGCAGGCG    14940

CGGTCACCTT CTGCAAAACG TGCCCAGTAG GTTCCGTCCG TCGCTCTCCA TAGAGCTCTC    15000

TCCATGAGGC GCGGGGTGCA CGCGCGGCAA GCGGTCTCAG GCGACTGGGC CGTATACCAA    15060

AAGACGCGsA TCCGTTTTTC TGCkTcGCAG GCAAAGCGGG GTCCCCACCC GTCATCTGCt    15120

TACACAGCAG TGTATCGTGC GTTCCCTGGT GTGCGTATAG CACCTGTTTG TCTGCCACGT    15180

AGCTTTCCAT ATCCGCATAC CAGTTTTGAT GGTCAGCCAT AATGGGAGTC ATGATGGCAA    15240

TCTCCGGGCG CAGCAGACCG GCGTGGTGTA CAGTGTGGTC CTGTGCATCG ACTGCGCGTA    15300

GGTCTGCAAG CTGCCAGCTC GACAgTTCCA GAACCACTGG TGTTGCAGGC GTTGTGTGAC    15360

GCACAAATTC CAGCGGGCTG ACTGTGCTAT TCCCCCCTAG AAAGGCGGGG AAACCCAGCG    15420

CACGCAAGCT GTAGCACAGG GCGCTGGCAG TGGAGGATTT TCCCTTGtGC CGCTTACTGC    15480

TAGCAGCGGG GCGGGAGAAA GGCGTAGGAA AAGGGAGATA TCCGTTTCgA tGgCGCGCCG    15540

GCGCkTTGAG CAGCGGAAAG GTAGATGTTG TGTGCACCCT TCACGATGGG ATTTTTGATG    15600

ACAACATGCG CGTTTTCAAA ATCTTCCAGC CGGTGTTCAC CGAGCGTAAA GCGGATGGAC    15660

GGGTACGCAC GAAgTCTTTT CAGGGAAGGG GTAAGCGCAT CAGCATTTCG CAGGTCGGTA    15720

ACCGTAAGgC GCgCTCCCGC TTCTGCACAA AAGnCAGsTG CCGCGCAgcC CCCGCCGTGC    15780

ACGCCGAGGC CCATGATGGT TACCGTTTTG CCTTGAAGAA GTGCGCGCGC CTGCTCCACG    15840

ATGCGGCCGA TTGTAGCCcG CGCAACGCGT GACAATAcAA GAGACGCGTG CGGTGCTCGC    15900

GGCACGGTTT CTCTTTACTT TTTGTTGCTT TTTTACTACC CTCGCGCGCT ATCTGCTTAT    15960
```

```
GGCTGAACAT ACTTCCTGTA CGAGCATTCA TCCTCTTGTG CGCAGCGCGT TTTACGCCGG          16020

GGGTGCGCAT GCAGTACTGC TTATTCATGG GTACATGGGC ACCCCGCGCG AGATGCAGTT          16080

TTTAGGTCGT GCGCTCCACC GGGACGGCTT TACGGTCTCT ATTCCCCGTT TACCTGGTCA          16140

CGGTACGAAT AGAGAGGATT TTCTTGAGAC CGGGTGGAGG GATTGGCTGC GGCGCGTGTG          16200

TGATGAGTAC CGTGACCTTT CCGCTGCGTa CCtTCGGTAT CTGTGGGGGG GCTGTCCATG          16260

GGAGGTGTGC TGACTGCACT CGTGGCGGCG CGTTTTTGTC CCCAGAAAGC TTTCTTTTGT          16320

GCACCGGGTT TTGCAGTTTC TGATTGGAGG ATAAAGCTGT CTCCTCTAGT CAGGTGGTTT          16380

GTGCGTGAGT TTGCTGCGGA CGCGGCTCCC TTCTACCCCG AGCAAGACTT TAATGACGCC.         16440

ACAAAGGATT ACCGGAGTGC GCACTACATT GCCCAGGTGG CGCAGTTTTA CGCACTGCAA          16500

AGACGTGCGA TCCGTTCGCT GGCGTGCATT CGGAGTACGT TGTTAACGAT CCTGTCTCGG          16560

CAGGACCCAT TGGTGCCGTG TGCAGCGGTG CAAAAATTAC TCGATGCGCG TGTGCGCACG          16620

CACACCAGTA CGTATG                                                          16636
```

(2) INFORMATION FOR SEQ ID NO: 48:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13330 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

```
TGATAAGCCC AGCGAATTAA TAACAAATCC TAAAAAAAGC GTGwArCCGA AAAAGGCCAT            60

AAGCGTGTGC AGAGAGCCAT GAGAAGGCAT TAGGTGCAAG ATGTGATCGC GTGGTACTCC           120

ACCACATAGC GCAGTAACGC TGCGAATCAA CGGAGCGAGG AGCGTACCAT GAGCGAAAAA           180

CTCCCCCGTC AGAAAACCCA TCACCATGCT AGAGAAAcCT ACGGACAAGA ACACGTAGTC           240

AAGGTGTGCC CATCTATTTA GCGCACGTAC ACGCCGCGTC CGCAGTAGCA AACCAAGTAC           300

GAAAAGAGG AGACCCTGCC CGAGGTCCCC AAACATAATC CCAAAAAGCA GCGCATAGGA           360

GAAAGCAACG AAAGGAGTCG GATCGACGAG CCCGTAGGGG GGACAACCAT AACTAGACAC           420

CATACGCTCG TAACTACGCA CAAAACGGCC ATGCTGGTAA CACACCGGCA CATGCTCGCT           480

GCCATCCCTG ATAAAAGACA GCTCCTGTGG TTCAAACAGG CGGACTGCCA TCCTCCCTGT           540

GGTCACGTTG TCCAATCCTG CAACGAGGTC CTTCGCCTCA TGTGCTGGCA ACCAGCCAGC           600

TATACGATAG GTATGCCGGG TAGACTCAAG CGCATCACGC GTGcGGTGcA CACGTTCCTG           660
```

```
CAGCGCAAAA CGCCTAAGCA GCGCACACAG TGCCGGGcCT CGCATCGAGC CGAGCGCACA    720

TTTCTCACCC TGaAACTGCT CTTCGTTGGG CAGAATCATG GGCTACAGAA CTCTTCCCAC    780

GGGGAACTGA GGGAGCATCA CCCTCTCGTT CGCGTGCGTG CAAAGCGTcA ACTTCTGCAA    840

GTATCTGCTG TGCCAACGTG TAGTCTTCTT CGGTAGGCAA AGAATCCCCT GGAGAAAACG    900

CGCACTCGCC AGAGACACCC AGGTACTTAC ACGCTTCTTC AAGACGGCCG ACATACTCTT    960

TGCTCTGCGC ACAGTGGGAA GAACTGCCAC GCGCCGCGGC GGAAAgACGC AAATGGACAA    1020

GCGCCGTCTT TCCCAGGTAC TCCAGTACTC GGTCCACATC CCGTTCAAGC ACTACAAGTT    1080

CGAGGAACTT CATCCTCTGC GATCTAAACA TAAGCTGCTT CCCCCAAGAA CTCACGTGGC    1140

AGCGTACGCG CACTCCCTGc ATGTATACAT TCTGCAAGCG CACAGATGCC aCGCGCTTCA    1200

AAGCGCTTCA TCATAAAAAA TGgACGATCG TCGcAGGAGT AAAGwrtGCG CATGAAAAAC    1260

CGTGCGCGCA AGACAATACA GAAAACGGGA GCCCGCACGT TCAAACGCGT CTAGATCCGG    1320

CTCCCACTGT ACCCCATCTA CAGGATTATT TAAGAAAGCG CTCATAGCGC CATCCACTCC    1380

ACGCCCCGCG CTCAGTAAGA GAACAATCCC ACGCAAACAG GATGCATTTC CACCACGTCC    1440

CCAGACACCT GCGTATCGCA CCGTTAGAAG AAAGcTCCGA AGGwrCnTyT nTCTTCTGTA    1500

CACCATAGAA CATGCGCAGA CGAAtACCCA CCCAACAGCG TACAACGCCA CTTCTTCACG    1560

CACTAACCGG GAACATGCAG TGCGATCATC TTCGGGAATA GAACACAaTG TCTGCCACAA    1620

CGCATGGTAG TAACTTTGAT CTAAGCGCAA ATCCCACAAG ACACGCTCAT CACTCCGAGG    1680

TACGCGGTTG TACCAGGAAA CTGGGCTCCC CCGAGTGACA GCCTCGATAC ACGGCCACTT    1740

TTCCCAACGA AAAAGGGAAA AACGACCCAA ATCTGGCGGT TGGCAGACGC GAAACCCCGC    1800

CTGACTCTCA GAAGAAGAGA GCGTCTTCAG GTAGAGATAG TCGTAACGAG CAAGCAGCGC    1860

GGAAAAAAGC GAAGGGGGTT CCTGATAGCA CGATGCCGCA CGCACACAAT CGCGCAGCAG    1920

ACACGCCTCC ACCCGCCGCT GGATCAGACC TACCAGATGT GCACCAGCAG ACCTAGGGGC    1980

AGGCTCAGAA AAAAGGCGCA CCCAGAGATC ACACAGCGTC CCCACACCCT GCAGGGTACC    2040

CAGGCGCTCA CCAACCCAAA GCCGAGCACA CAAACCGGAT AGTTTCGCAT AGACGAACAC    2100

ATCCGCGACA TCTCGCTCCA CAGCCACACA CTACCCAAAA AAGAAACCAT CAAGGAAGGC    2160

TTCGAAGCGC TCCACGTCCT CCTCCAGCAC AGAAAGCTCG CGTTCTAACT GTCCACACGC    2220

CTCCCTTCTC TGAGCGTCTA TCCGCGCACA GGCCTCATGC CACGCCTGCT CCTGCGCCTC    2280

AACCACGCTC CCCTCCGCTT CCTTCACCAA GGTCTCCCCT GCCGCCTCTG CAGATTCGCG    2340

CAtCCtGCGC CTGCTCttGC GCCTTCTCCA CGAGGACAGC GGCCGCATCC TCCACCTCCG    2400
```

```
CTAGGCGGAC CATCACCCCT TTATCCATAT TACCAACCCA CGCGGCCGAC ACTCAGCGAG    2460

CACATCGACC TACTCCCGTA CCTCAGGGGA CGGAGGAACC GCTACCCCAG AGCCCGAGGA    2520

ATTCCTGCGG CAGGACCAGT AGACTCAGGA GACGTTGGAG CACCAGCACC TTCTTcGGAG    2580

GkTGCTTCCA CCCACTCCACa GCcGTCTCGG CTTGCTTcTg TGCGCCGCTT TCTGCAAGCC    2640

CGTGTCGTCC GGAGCCCGGT TCAACAACGC GAGAAAGAAA GTCAAACCAA AAAAGAGCCC    2700

TACCATAACG TAGGAAGTCT TCGTGAGGAC CGAAGCGGAG CGCGAACCAA AGGCAGAACG    2760

CGAACCGCCC GAAAACATGC CACCAgCCCA TCTCCCTCTT CAGTCTGCAA GAGsATAGCG    2820

TGACCACCAG AgGCAAACCA CCACCAGGAG TGAGAGTATC ATAACGCTCA GCACAGCCAT    2880

GCCTCATGCT ACACCAAAGC GTGCGCAAAA ACAAGGACTT CAACCTTTCA CTCCCTGTCC    2940

CGAGAACAGA AAAACGACCA TGTGAACGGC ACAAAAAAG AAACCTGtTA CGCAATACCG    3000

CACACCACCC AGAATCCTTT ACTACTCTCT ACACtGCGCG CGATAGGAAC AAAAGACGCA    3060

GCCTCCAGCG AAcACCGCCA ATGAGTCCCC CGTCAATGTG CTCTTCAGCC AACAGTGCCC    3120

GCGCGTTCTC CGCTTTCATG GATCCGCCGT ATTGAATACA CAGTGCCTCT GCGATAGCCG    3180

CGCCGTACAT CTCGCGGACT ACTGACCGAA TATGAGCATG AACCGCATTC GCCTGTGCCG    3240

GAGTGGCAGT CTTACCCGTA CCAATTGCCC ACACAGGCTC ATACGCAACA GTTACATTAT    3300

GCATGAGTGA CCCACACACG TCTGCCATCC CTGCGCGCAC TTGAGTTCCC ACTACCTCGT    3360

TGGTACACCC CGCTTCATAC TCTTGGAGTC GTTCGCCGAC GCATAAGATG ACGCGCAAAC    3420

CGCTTTCTAA CACGCGTCTG ACCTTTTGAT TGATAAGCTT ATCATTCTCC CCACGCCCAT    3480

GACGCCGTTC GGAATGCCCC ACGATGACTA CCTGTACCCC CAGGTCTTCG AGTTGAAGGA    3540

CGGATACCTC TCCAGTATGC GCCCCCCACT CTTCACTACT CACGTCCTGC GCGCCAAGAA    3600

GTACGTTACT TCCCCGTAGC ACCTTCCCCA CCGCGTCTAA AGCGGTAAAA CTCGGCGCAA    3660

TCATGTATGT GTGCGGACCA CCCCGTAATT CCCGCACGAG TTCCTGCGcA AGGcCACCGC    3720

CTCCGCACAC GTTTTATGgC ATCTTCCAAT TCCCCGCGAT AAAATAGCCG CGCATATCCC    3780

CTCCTTAGCA CATcCTTTCG TTCACACCAC AAACACCCCG CCGATAGCTC CACGAGAAAC    3840

TGTTCTATAG AGCCACCGCA GAGACATACC CTCGCCAGGA GATCTCGCCC CAACCCAAGG    3900

GAAAATCTCA CGGCACCACT ATATCCTATG TTTCAAGACA CGAAATGCCC GGTAAAACTT    3960

TACCCTCAAA GAGCTTCAGC GATGCGCCAC CCCCCGtAGA TACATGGCTC ATGCGACTTG    4020

CAAGCCCAAA CTTGCTGACT GCTGCAATAG AGTCTCCTCC ACCAACTACC GACGTAGCAC    4080

CCGCATCCGT CGCCTCTGCT ATCAACTGCG CAAsACCCGT GTACCGTGTG CAAAGGCATC    4140
```

```
GAACTCAAAA ACCCCTACCG GACCATTCCA CAACACGGAG CTCACTCCCT TTAAATGCGC    4200

ACGATACTGC TCAAGCGTAC GCGGACCAAC ATCCATACCC ATCAAGTGCA TAGGAATATG    4260

CACATCGTCC ACCGcAACCG GcTGCGCATC CGCACAGAAC GTGGaCGcAC ATaCGTGaTC    4320

GACCGGCAAT aCCACCGaCA CACCACCGcT TTGaGCCTTT TGCAACAGCA TaCGTGcaGT    4380

GTCGATAAAG TCATCCTCCA CTAGGGAGGT ACCTACACCC ACACCTTGCG CTTTCAAAAA    4440

GGTGTATGCC ATCCCTCCCC CGATGATAAG CGCCGTCGAT GTTCGAAGCA GACTCTCCAA    4500

GACTGCTATC TTAGAAGATA CCTTGGCACC ACCGACAACC GCCACCATTG GCACCTTCGG    4560

GTTGCATACC ATAGGTTCCA GGTACCTCAC TTCCCGCTCT ATCAAAAGAC CGGCCACTCT    4620

GCGACGCATA AGTCTCGGGA GTACCACCGT AGATGCATGT TCACGGTGCG CAgTGCCGAA    4680

CGCATCATTA ACAAAAATGT CCCCATACTG GGCAAGCTCC CGCGCAAATT GCTCCTGCAC    4740

CTTCGCATCA CCAGATGTTT CCTCGGGGTG AAAGCGCACA TTCTCCAAAA GCACTACCGA    4800

ACsGTCGGGC AGCCCTTCAA TAAATTCACG CTGCCCGACG CAGGAAGGCG CAAAATGCAC    4860

CGGCACCCCC AACTTCTTTG CAAGGCAGTC CGCAACCGGC TTAAGCCGgT GTTTGCCGTT    4920

AATGAATGCG TGGCGGTCAA AGGGACAACC ATCTTTCTTA GCGTTCCCCT CTGCTTTATC    4980

CGCATCACGG GTAGGGTCTC CAAGATGGCT AATGaGCACT ACGTGTCGCG GCCCCTGCTC    5040

GATGATGTAC CGCAGAGTAG GAACTGCTGC AGTGACGCGC GTGTCGTCTT GCACCATACC    5100

ATCACGCATC GGTACGTTAA AATCAACACG CACGACAACA CGCTCACCTC GCATTGTGAC    5160

ATCTTTACAA GTTCTCAGCA TCATCTCCTC CTTTTTGACG CAGGGGTTAC CCCATCCGCC    5220

ACACGTTGCG GCTGATTCTC ACTATATTTC AAAAAAAGAT TCAATATCCG CATCGGTCGG    5280

ACGCCTCACA TCACTGCTCC CCTGCCTGTG CGCCCTACAC GCGTACGGGG TGGGGCACAG    5340

ACCCTTCGTC ATTGTTGACA TTTTCTATGC GGAATGATAT ACCCCGGCGG GTGCTGATTC    5400

CCCTGTAGCT CAGTTGGTAG AGCAAATGGC TGTTAACCAT TGGGTCCGTG GTTCGAGCCC    5460

GCGCGGGGGA GTGATGTTTT TGGTTCTTTC AGTTAAGAAT TCTCATGGAA GGTGGTGTGT    5520

CTTTCACGGG TTGTGGCCCC TTGGGGGCAG TGAGCAGTAC TTCCAGCTTT TTTAGAATGG    5580

GCCGTGCAGC GTCCCGTGGG TCTGTGTGCG CCTGGTTTCT ATCGGAAAAA TGCGGGGCTT    5640

GGTGCCCATG AGAGTAGGTG CCTATGGAGT TGAAGGTCCG TCAGAGTGGC GGAATATGTG    5700

TCGTAGAaTC AGTGGGGACA TGGATCTGTA TCATTCCTAC AAGCTTAAAG ACCTTGTGCT    5760

GAAGTTGTTC GATAGGGGCC CGCGCTGTAT CGTCATTGAC CTTGAGGCGG TAGAGTATAT    5820

CGATTCCTCA GGGATTGGCG TTCTCATCTA TCTGTGTTCG ACAGTGAAAA AGTTAAAAAT    5880
```

467

```
CCACTTCTTT ATCTCAGGTG TGCACGGCTC TGTAAAGAAA GTGATTGAGC TCACCCGGCT    5940

GCTGAATTAT TTTCCCATCG CTGAAAGygT AGACGAGGCT CTTGCAAGGG CCCGATCCTC    6000

TGCACCGCCG CAGACCGGCT CCCTGTAGGT TTTTCCTCGT CATGGGTTGA ACCCTCCCAC    6060

GCGCGGGAGG GTTAGATATC CCACAGCTTT TTGCTGCCGc GTGCCGCTGC ACGGACTGCA    6120

GTGGTAGCGT CTTTCCTTTA TACTCAGCAC TGTGCATATG GTAACGGACG GCGCTTCCCC    6180

AAGAAGTGGG GTGTCGCTCA TTATCGGCAG ACCTTCCTCA GGTAAGTCAA CCTTTCTCAA    6240

TGCCGTGTGC GGGTACAAGG TGTCCATAGT TTCCCCTATA CCTCAGACAA CCCGTAACAC    6300

GGTGCGCGGC ATCGTAAATA TAGAATCCGA CCAAATTGTC TTTATGGACA CCCCGGGGTA    6360

TCACCGGTCT GACAGAAAAT TTAATCTGCG CCTGCAGTCC CTTGTGCACA GTAATGTAAA    6420

GGATGCTGAT GTGCTGTTGT ACCTAGTAGA CGCTACCCGT CAATTTGGAG AAGAAGAAGC    6480

AGCCATCTGT GCATTGCTTG CCCCGTATCA AAAAACGCGC GTATTGCTTG CCTTCAATAA    6540

AGTGGATGTC CTTCACAATT CGACCTCGTG CGACGAGCAT GCCTTTTTAC ACAGGCAAGG    6600

CAGCGTGCTG CGGGCCGGCA GCCTGGGACG AgCGCTACAC GCCGCACTCC CCCACCTCCC    6660

TGCTGATCGG GTATTTACAA TATCTGCCCT GCACCAGGTT GGGCTCGATG CCCTCATGCG    6720

CACGCTGAGA GATCTCTTGC CAGAAGCGGC GCCTCTGTAC CCTCAGGATT GCTATACGGA    6780

TCAGACCATC GCCTTTCGCG TCACTGAGCT CATCCGAGAA CAGGCAATCG CACGCTGCCG    6840

GGACGAACTG CCGCACGCAC TATACGCCGG AGTGGAAGAC ATGGAGctGC GCCGCGGCAA    6900

GCGGGAACTG TGGTGCCGTG CGTTTCTTGC AGTAGAACGG GAAAGTCAAA AGGCAGTGCT    6960

CGTGGGGAAG AAAGGTGCAG TTATTCGCgc CATACGGCTA GATGCCATCC GCGCGcTACG    7020

CACACTCCTC CCCTACCATA TTTCCCTTGA TATACGAGTG AAGGTAGACC GCAGCTGGAG    7080

ACAACGCGAC CACACACTCA GCTCCCTTCT GTACTAGGAT GACCGGTGCC CAAATGAGGA    7140

ATTCGCCGCA GGGGCGGGCC GCTCAAGGCG TATAGTTACT GAAGGTTCGT CACACACAGC    7200

CGGAGgTCCA TAATACTGTA CCGCCCCCGG ATACACGTAG CTTGTTTTTA AGGCCCAACT    7260

CGCACGCCGA CGAGAGAACA CCCGAAACGG CATGCCCTCC AGGTCAACCA GCGCCTTTTT    7320

TATCACCGGC TTTTGACTTC CATGTCGGCG CTCCATGTTC ATGAGCATAG TAAGCGGCAC    7380

GCCACCTACT GCCCACTCAG CTACAGAAGA CGTCAAGTTA CGCACCGACG CTACGTACCC    7440

TGTAAACCTG TGCACGGCCA GAAGACACGC GGTCAAACCG AGCGTATAGC AATAATCTGC    7500

ATCAAAGTTG GACGGAAAAG CGCATCGCCC TTCGTAACCA AAAAAATGAG CAATGCTGGA    7560

AAAAACACCG GTGTACGTAC CTTCCTGCTT CATCTGCGCT AAGCGCTCCG TTACCTGGAG    7620
```

```
AATGAGCAAA CGCTCTGTGT CAATGCGCGA CACCTGCACA TTCCCATGTG GATCCCGATC    7680

TGCCAAAAGC TGTGTGGAAA TTTCAGCAGG TAATGCGTTA AACACCGCAC GAGCAGAAGC    7740

AGACAACGCC TGCTCTATCC AAACGCGCTG CGGTtCAGGA GTGTCCAGCG CCTCAAATTC    7800

CTGCGCGCGG CGTGCCATCA CCTCGTTGAG CTCCGTAATT AGAGCCTTCA TTTCAGGTAT    7860

AAATTCGATA AGTCCTTCTG GAACTAACAC TATACCAAAG TGCTCACCGT GTTGTGCGCG    7920

CGTGGCGATG GTGTCACACA ACGACTGCAC GATCTGTGCG AGCGTTAACG ATTGCGCCGC    7980

TACTTCTTCC GAAATGAGAC AGACATTTGG CTGTGTTTTC AGCGCGCACT CAAGCGCAAT    8040

ATGACTGGCT GAACGCCCCA TGAGCTTAAT AAAATGCCAG TACTTGCGGG CACTGCACGC    8100

ATCGCGCGCA ATGTTCCCGA TAAGTTCACT GTATGTTTTT GTGGCAGTGT CAAAACCAAA    8160

CGAGGTTTCT ATCGCCTCAT TTTTCAAGTC TCCGTCAATA GTTTTGGGAA CACCGATAAC    8220

CTTGGTAGAA ATACCACTGT TTACGAATGT TCTGCCAAAA GGGCAGCGTT CGTGTTGGAG    8280

TCATCACCTC CTACAACTAC GAGTGCATCA AGCGCCATAC GCGTGACTGT CTGCGCCGCG    8340

GmGGcAAACT GGGACTCACT TTCGATTTTG GTGCGTCCTG AACCAATGAG GTCAAAGCCA    8400

CCTGTGTTGC GGTAgcATtC TACACGGTCT GCGCATATCT CGATATGATC GCCAGAAAGC    8460

ACGCCCGCAG GACCGCCTAG AAAACCGATA AGGACAGAGT CAGCGTGCCA TCGTTTTAAT    8520

CCGTCGAAAA GCCCTGCTAT AACGTTGTGA CCACCTGGTG CCTGACCCCC TGAGAGTACT    8580

ATGGCAACAC GTAATCCTCG CGGCTCAGGT GCAGTCTCCA TGGGGAATC TTCGTTTTTC     8640

TCACTAGCAT TAACGAAnTT CACCAGCGGC TGACCGTAGy CGGCGcAAAA AGAGAGCGCA    8700

ACGgTcATAG TCTGCCACCG CAgTGGTGGA TAAGCCGCGA CGCGCACAAA CGCGCCGAAA    8760

GTCCCCCCGA AGAAGATCGG GGACCTTTGG CAGGTAGCGA TGCCGTTCCT GTTGCAAGAG    8820

AGAAATACTC ATCGATGATT ACTCCTTCAT ATACGAAAAA TAGCACGACC GCACCGCCGC    8880

ACCCCCACAA CTCACTCTGC AGCAGGCGCG ACCGCGTGTG GATGCGCAaT ACTCAACGCA    8940

aGAaTAGCAC GTTTAAGAAC CGTCGCTTCT TCTTCATACA GTGAACGCCC CACACTGCAC    9000

AACGCAGCGC TCTCCTGTGC GATCACCTCC GCCGCCATTT TCCTGTCGTA TCCCATCTGT    9060

ACAAGAGCAG TTACCAGATC CTCAATTTCC CTCGCATGGG GAGCACACCC AAGATTGCTC    9120

GGATGTGCAG CACGATCATC TGTCTGACTC TGGGCACAAG AGGCCgCGTC GGTTAGCGCG    9180

AGCGTACCTT TCAGCGCTAA GAGCATGCGC TGTGCAGTCT TTTTTCCAAT GCCTGGTATG    9240

CGCTGGAGTG CACATAAATC TCCTGTATCA AGCGCTGCAC ACAAAGCCTG ACTGCTAATA    9300

CTCGAAAGAA CTTTGAGCGC CTGCTTTGGA CCAATACCTT CTACCTTTGT AAGACTGAGA    9360
```

```
AAGAGCGTGC GCTCTTGTAC ATTCGAAAAA CCAAAGAGGC GAAGCGCATC TTCACGGTGA   9420

TACAACCAGG TAAAGACCTT AACGTGTGAT CCAACCTCAC CGAACGCAGC ACTACTGTAT   9480

GCGGACACTG CAATTTCCCA TTCAATACCA TGCACCTCAA CACAGAGGCG CTCGCGCTCA   9540

TGGAGCGTCA AGATACCGCT GATGCTTTCG AACATTATCT CCTCTTTATG TGTGGCGCAC   9600

TGAAGCGAAC ACTACGCTAC CATGACAATT GCACAAAGAC CGGATCGTGA TCAGAAACGC   9660

GTTCCCGAGC AGGCTGCTCT GCGTTGATAT GCAGTATATC AGCAGTCTCC GTGCGCGCTC   9720

CTACAGACAG GATATTATCC AGTGTTTGCG AGTAACCGCG GTACACATAG GTATATCGCT   9780

CCGTCTCCGG CAAGAGATCC AACGCACTGT GCATCCCCAC TGCGGTGAAT TTTTGAATAA   9840

CATCAGAAAA CCAAAAATCA TTGAAATCTC CCGCCACCAC CACCGGAAGA TCTGCACGCT   9900

CGCGACGTAt GCAGCAACAA AAGCGGCAAC CTGCGCCGCC TGCTGTATAC GCTTGCGTTT   9960

GGAGTGTTCC TGTGCAGGTT GCGTGCTACC CCAAACGGGG TCATCCCCTC GCTTTGAAGA   10020

AAAGTGATTC GTTACCACAA CAAAATCTTT CCCCTTATTC ACCCTGATA CAAACTGAAA    10080

ATGTGCCACC AAAGATTTAC GTGTGTTTTG AAAACTTTCT TGCCCTACTC CGATGCGCGC   10140

AGGATTTTTT ACCATCTGTC TTCCCCGCG CACCATTTGG GCAACCGAAT GAAATGTTCC     10200

TGCACTTCCT GTCTGGTCCT GCACCAGCTG CACACGATCG GTAGGTACAA ATAACAACAG   10260

CGAATATTTC CTCCCGGTTG TCCGCCATCG GCATCCAACG ATTGCACGCC CGCaGGAGCa   10320

TTaTCCTGCG GATCGATATT CACCGCTTTA TACCGAACGG CGCTGAACTC TGCCATtGCa   10380

CGTACCAGTA AATCCAACGT GTGCTGTGCG CTCGTACAGT GATGATGTTT TTTTGCGCCA   10440

TCGTCATCCT GTATCTCAAC AAGACAAATA ACGTCCGGCG CCTTAAGATC ATTCACAAAG   10500

TnTTCGCAAA GACGCGCGCA CGCGCTGAGT CTGCTTTATT CCCTGCAGAA AAATTCTCCA   10560

CATTATAACT CGCTATATTC AAAAATCGTG CGTTGAACTG TATGGTCGAA ACTTCAGGAC   10620

TAAAACCTGA GCGTCTCAAA GGGGGAAGCG GCTCAGCAAG TTCTAATTGG TAACTAGAAG   10680

ACGAATACCC CATGATCCCC ACCACCGTCC CTTCAAAGGA ATCACCAGGC AGAGGAGGGG   10740

CGCTTTTGAA TACTTCAGGA AGGCTGTCAA ACATACGCCG GGGACAAAAG GCAAGGACAG   10800

GACGTATATG GGTTTGCTCA TACACGTATC CTCCGTGCAT ATTCAAACGT GTAGAAGGGG   10860

TATCCCCCGG TAGGAGGTAA TACGTAGAGC GATACGCAAC AGCAGGAACG GTGGGATTCA   10920

CCATCTGAAC CCGCATCCCT TCCACACTTT CATAAAAATC AATAGTCTCT GCATCCGGTG   10980

CGAGGTCTGC AAGGTTGCTG ACAAACACCG GCTGAGACAC CCGCGCATAC GAAATCAACA   11040

CCGGTTCAGG CAATTCCCTG CCATGTGCTA GCACTCGCAC ATCCTGCGCG CGCTTGATAA   11100
```

```
CAAGCTGGGT GACGCTCAGA TCCCGAGCAT TGCCTTTTGA GATATACTCG CTGACAGTAC    11160

CGAGCACCGC CACGTAGTCA CCCACGCGCA AACTATcAGG GAAAGCCTTA CCACAATACA    11220

CAAAAATGCC GTCAGACGTT TTAGGATTGC CATCCCCATG CGGATCTTGA AAATAAAAAC    11280

CAATAGGTCG TTTACCCGAA CGCGCAATAG CAGTTACCAC GCCACGCACA TCACGCACGT    11340

GTTTACCCTC ATAGGCAGAA CGGTGTCCTT CCCCTTGGAT CGCACCGATT GAGTGGGGAA    11400

CAGACGCCGC ACTGCACCGT GATCCTGTTC CCACTATCCA AAAGATGACC CCCACGCACG    11460

CTCCCGCTAC TTTACTGCTC ATAGGACACT CCTCACGCGC AGTGTATCAG CGCAGGTAAT    11520

TTTGCACAGT ACGGTAGTTT TCTTTGGTGA CAACTTTATA GGGGATCCAC ACACACTGCT    11580

TTTCTGCCGT ACCGAATACT GCCGCGCGTC CTGCCACTCC GTTTCCTGCG TGCGTAAACG    11640

CAGTGCTGTC TTCAGCGGGG ATCGATGGGA CAGAATCACC CATAAGTGCA AACAAAGAT     11700

TTAAAATAGC CTTCCCCTGA CTGGAGACAT CGTTAAGGAC GGTGCCGAGC ATCAGATCCT    11760

CTTCAATAGC TTTCAAAGCA GACGCAtAGC ATCGATACCC ACAACCGGCA CACGCTTATT    11820

TTCTTTAAAA AAACCTGcAC TCTGCAACGC TTCAATGGCG CCGAGCGCTG CGTCATCGTT    11880

ATTCGCAAAT ACTGCCCTCA ATGCGATCTC CGTGTGTGTG AATAAGCGTG TGCATCGcAG    11940

CyTGTCCTTT CACCcGACTG TCAaGCGCAA AaGCCTCCCc GATTATCTCG CCcTTTAATC    12000

CGATTTCTCT CAGCGCCTGA CACACATACC GCGCACAGCG AGCACCGCTT TTATGATCAG    12060

GATCCCCTTT GAGCAcTACG CATTGGATAA TACCGTCGGC GTTCTTATCT GCACTTGGTG    12120

TACGTTCCAG ATATTGCGCA ACCAGTCTGC TTTGCAGCAA ACCAAGCTCG TCGTCCTTGA    12180

CGCCTACGTA ATAGGCGCGT GCATACCGGT TCAAATCAGA AAGGTCAGGC ATACGATTGA    12240

AGAATACTAG CGGAATGCGC GCCTGCTGTG CCTTTTCAAT AACCGTGCGC GCAgcaCGAt    12300

GGTCTACAAG ATTTACCGCA AGACCGTGCA CGCCGCGCGC AATAAATTGA TCGATGTGCT    12360

TGTTCTGAAT ACTCTGCGAT GCCTGACTAT CCACGATGAG GATTCGAGCA TGTTTTTTGC    12420

CAACCGTAGA GAGTATGTGA CGCAAGCGCG CCACGAGCGT GTTGTCATAC TGATACACGA    12480

CTACTCCGAT AGTCGGCTTT TCGCTGCGCT TGCACGCGCC CGCACCAAGT GCACACAAAA    12540

GGAGCGCTAC ACACATCCCT GTACCTTTCA TATTTCCTCC TCATGTTCAC CAGCGCATTC    12600

TGATTTGACA CTTCTTTCCC CTCACACCCT GATACCCGCG CGAGGAATAT AGAAATTAGA    12660

AAAAGGATGG ATTATCCAGT GCTGCCACCA ATCGCATGAA CGTGTCTATG TACCCGGCCT    12720

TGCGCCGTTT AGCGTACACg TCTGCGACAA TCGCCTCACT TGCCTCAATC ATGTATATGT    12780

TTTGGATTTC TATTACGTCG TACGTGCTAA AGCCATGCTC ATCTGCAAAA TTGCCAGGAA    12840
```

CCAGTGACTC CACACGGTAG GTGTTGCGCG TGCCGGTGCc gCCAAGCGCA TCCCAAAAAG 12900

GGGCAAGAAG GCCCGGTACG ATAAGTCGTA CTTATACAGG ATCCGGGCAG GATGTTCCGG 12960

CCGTACTCCT AACTGGCAGT ACCATTCATC CTGCTCATAG GCGGCCAGAT CATGCAGTTT 13020

GCTCTTGTCC CGCAGTCGGT ACCCTGCAAG ACGGACGATA GTTTCGGGCA CGTATTTTTC 13080

CAACTCTGCC TGTAAGTCAG CTACACAAGA AACGGATGCA TCATTGACCG TGGTAATAAT 13140

TGCCCCTTCT GGTATTCCTG CGTACGAGAG AGGACTGCCA GGAATAACGT ACGAAGAAAG 13200

CACACCGCCG ACGCCAGCAT TTTTCCACAC ACGGTGTGTT TCACCAAACG CACCAAGCCA 13260

CGGATGCGTC ACCAATCCCC CGCGGTACAA GTTGGCAGCA CCTGCTTGAG CAATTCTACA 13320

GGAAATGGCA 13330

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 10214 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: double
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

ACACGGTGGC GCGCAGATTA CGAAGAGGTT AAGCAGCTCG GTGGTTTGTA CGTCATTGGC 60

ACAGAGCGGC ATGAAAGCAG GCGCATTGAT AACCAACTTC GGGGGCGTTC GGGGCGTCAA 120

GGGGATCCAG GCCGCTCAAA ATTTTTTCTC TCTCTGGATG ATGATCTTAT GCGCATTTTT 180

GGGGGGGAGC GGCTGAAGCG TTTTATGAGC CGTGTGGGTA TGGAACCAGG AGAACCTATC 240

ACGCATTCCT GGTTGAATAA GAGTATTGAG CGCGCGCAGA CGAAGGTCGA AGCACGCAAC 300

TTTGATGTCC GTAAGCACTT GCTGAATACG ATGATGTGCT CAACGAACAG CGCTCCTTCA 360

TATACgCGCA GaGcACAAAT TTTGATAGAC GAGCATGTGG TAGAGCGCGT GTATACCACA 420

ATCGAGGAGT ATCTTAACCG AGAAATAACC GCACTTCGGC AAGAATTGAA GCGGCGTGGG 480

cGGCTTTCCC TCGGGGCGTT TCAACAAAAC CTGAGCACCC TGTTCGATTA CGCACTGGGA 540

GGTGAGGACG CATCTGGCTG GAACGAAACG CGTCTTGGAA CGCTGAAGCA AGAAATCCTG 600

GCGCATTTAA AAAAGAATAT TGAATCAAAG TATCTGCTTG CAGGGGCGCA GAACATGGAT 660

ACGTTCATCC GCTACCAGTA TGTGCAGGCG ATCGATAAAA AATGGCTGGA CCATTTGGAA 720

CTTCTTGAAA TCCTCCGGGA ATCGGTGTAC TTGCGTTCAT ATGGGCAAAA GAACCCGCTT 780

ACCGAATACA AGCTTGAAGG GTTCGACCTA TTTTACACCA TGTTAGACGA CATTCGCCTT 840

```
TCGATCGCCT CGCAGGTTGT GCGCGTAACG GTTCACATGG AAGAGCAGCG CGTCCCGAGG   900

CCACCACACt TGCACAGGCG GCACACGAAT TTCAAGCACT GGGGCAGCCT GGCAGAGGGC   960

ACGGATCGCT ATCTGCTCTC CCGATTCAAG CCGGCGCAAA AGTGGGGCGC AACACCCCcT  1020

GCCCcTGTGG AAGTGGCAAA AAGTACAAAC ACTGTTGTGG CCGCTGAAGA GCAATCTCAT  1080

TATTTTGCTT GATGGGCAGG ACCATCCAGA TGTCTATCCT GTTCCAGGTA AAGACCGCCG  1140

CTCAGAACAG AATGATAAAT TCTTCAAGAA AGACATGGGT AACTCTCCCG AGACTCAGCT  1200

GTGTGTTGaG CGCATCGCcA AGGCAATTTT TAGCTGCTCG GGGCCAAGCA TATGCAATTC  1260

TTGCGCGGTG TGCTGTGCAC aaCGCAAGAG ACGCACGTCT CAGTTGTTCT TTTTTCTGAA  1320

CGAGCTCCTC ATACAGGGCG TGATCTGCGG CAGGATATTG CAAAAGTGGG GTAATGATGA  1380

CACTCACAGG CTGCTTATCC GCACTCAGCG CGCGCAAGTn CCAAGTTCTG CATAGACTGC  1440

ATGGGCAGAT TCCCGCTGGG CAATCCCCCG GGTTGTACGC CTAGTCTGCG TTTCCCTTGG  1500

AGCACTGTGT GTCCTCACAC GGAACGCCCC GCAGTGCCGA GAAGTAACAC ACAGACGATG  1560

AGCGCTGCGA CAGTTCTCAA TGCACGGATA ACACGTTGTG CAGTCTCCTC AGTCATGGGG  1620

CATTGTAGCA CGCACAACAC TCACTGCACA GCGATAAAGA CTTgCTTGAC AGCACCCTTG  1680

TACCCTCGTA CACTGGGGGC GGGCATGGGT GTTCTTTCGT GAAGACAAGT CTGTTGCTTT  1740

CCGTTTGCGC mgsGCTGCGC TGTCCGGTTG TGCCACGGGT CAGAGTGATG CGGTCACAGA  1800

CCCGCTCTCG GTTCTGGAGG TTTCTCAGAC AGAGACGAGA GAGGCGCTGA TGCTATTTGT  1860

CTCTTACAAC GAGACGGGTG CATCTGTCAC CATCTTTACC CCTGAATTGG TTGCGCGTCT  1920

TTCCAAATCG TATCGCTTTC TTCGCGTCGA GGCTCCTCAC AGCGCATACA CCCTTTCCCC  1980

TGAGGCGCGC GAACGTAATC GCTTGTTGTT TTCGGAGTAT GAGGTTGATG GCCTTCCGTT  2040

CCTTGTTCTC CaAAGCGCAC AAGGGGACGC TTACTTTGCG CAGCGCATAC ATTCGACGCT  2100

GTCGAGCGAG CAGGAGCTGT GGGCGCTAAT ACGGTCTGCG GACGCTTCGA GAAAAAAAGT  2160

GCTGGCGGCG CGTGACCGTA TCGCTCAGAC CGAAGCTGCT GAAAAAGCAA TTGCCATCGA  2220

TGCATTTCTT AAGACGGTGC GTTACCCACG CTCTGCGCGG TACGACGCCC TCCGAAAAGA  2280

AGCACTCCAG GCTGATCACG AAAATGTCTC AGGTCTCCAC GGGGATTACA TGTTTCACCT  2340

GGCACGGCGG CGCGCAGAGA AATTTATCAA GCAAGAAAC CTTGTAGCAG CGGGGAATGC  2400

TTACAAGGAT TTAGCGCAGT CACCGTTTCT GAGTGCATCT CAAAAACAGG AAGCGTGGTA  2460

CCTGACCGCA TACACCTATG CTCTTTCAGA AAAGGTATCT ACAGAGGACG TATCGCGTGC  2520

TTGCGAAAAG CTGTTGCAGC CCATCCGCAT GCTGCGCGGG TTGCACAGAT CAAGCAAACC  2580
```

```
ATAAAGAAAC TACTTACCGA GAGAGGcATA TGAACGAGGT CGATGCAGTA AAAAAGGGAT   2640

AGACAGTGcA AATCCTGCAG CAAGTTGCTC AGGGGATCGC GTCTCATTTC GGGCATGACT   2700

GTGAGGTGGC TGTGTACGGC GTCAGTAGCG ATGGTAAAAA CTGCGCGGTT GATTTTATCA   2760

CAAATGGACG CGTTACCAGT AGCAGGGTTG GAGACAGACC CCGCCTGTCG CTCTTCAAGA   2820

ATTACGGAAT AGAAACAGGc AAGGGCGGCT CAACTACCTC ATtCGCACGG AGAACTGCCG   2880

CTCCCTTAAG TCGAGCATGT TGTATATTCG TGACGAACAT AGCACGGCTC AGGCGATTCT   2940

AGCGATAAAC TTTGATATTA CTGCTTTGTA GGTTACGCAt TTGCGCTTGG CCGGCTCACC   3000

GGCACTGCTG CGGAGACCGC CTCGCATATC CACCTTAAGA GCGTCAGTGC GTTCCTCGAC   3060

GACCTGATAG AAGAGTCTGT AGAAAGAGTA GGAAAACCTG CAGCGCTCAT GAGTAAAAAG   3120

GAAAAAACGG ATGCCATCCA CTTTCTCAGC CAGATAGGGG CGTTTCTCAT TACACGCGCG   3180

GAAGACAGGG TCTCCCACTA CTTCGGCATT TCAAAGTACA CCCCTACAGT TATATCGAAA   3240

CTGGCAAATC GTGATCGCAC CGGACTGAGT CCCCAGCAGA GGGATCGCCG GGCCCTACTC   3300

CTTCCCTGGT TCAAGCTCCT CGGcGAAGAC AACTCCTCCG GAGCGGACCG CTCGCACCAC   3360

GCTCCCACCC GTCCTGAAAT ACTCGGACAC CACCGGTGAG GTCGCCGCGC GGAAAGATTC   3420

AACTATTCTT TGTCCCACTT CGTCCCCGTT CCGCGGGATC CGCCGCACCC AATATAACGC   3480

AGCGTTACGC GGGTCTTTGG TAAGCGCGTG CACCTCCCCA TCCACTAGCA GCCTTTTAGG   3540

CACCCCCTGC ACAAAATCTA CTGTCACGTG CCTACCTGTC ACCGGGTGCA ACCACTCAGT   3600

ACGCGCATGA CCGGTAGGTA GTTCTGTGTG CGAAATCGTC ACACGTCCAC GGCCATACCA   3660

TTTTTGCACC TTCTGTCCCT TCGCTCCATA CTCTTCTTGA TAGGCAAAAC TTCGATCCTT   3720

GTGCACGTCA ACTGACAACG CACGTACCGC CCCTTGAGCA TTGTAGTATT CCCGCGTTTC   3780

AAAAAATCCA TCATCATCCC GATCGCTATC CCGTtCGCGT TGCGCGTCCA TCCACATAGT   3840

GCGTACGCGC ACGAAAACGC GATCCAACAT GAGTTTCAGA AAATAAAGGC AGCCCTTCAT   3900

CAAGGTACGT CCttTACGCGG GCACGCTCAA ACAGGGAATC AGGCTTTTCG TAATAAAGGG   3960

AAGAAACTGT GATCTGCTGC TCAGTAGGGA GCGGCTCATT TGTCAATACC ATTGTGAAAA   4020

AATCGTGCGA CCGCACACCT TCTAAATCTC GCGCAAGATC TAAGGACTGC ATGCGTACCG   4080

GCTGCCACCG AAGTGCACGA GGACGCAATA CGTACGTTTT ATCCTCCCAC CCCACCTGGT   4140

GTACTTCAGG GTAGCGATCG TAACACACTC TGTAGCCTTG CTGCGTCAAA GGAACACGCG   4200

CCTGAGTTTC TCCCTCCACC CCATTATCTG CAGGAAGCGA CACGCGCGGG GCAATCCAGC   4260

GCTCAGCAGC ACGCTCATGG TCGTGCGCGG CAATATTCTG GGGTATGGGG AACACAGAAG   4320
```

```
CGCGCGACAA TTCCGTCTTT TGCGTACTGT GCATCGTGTG CACGCACGTT GGCGCACCGT    4380

CATTCGCATA GACTTCATAC TCGAGGATGC CATCCTGGTT TGTATCGAAC TGTGCCCGGC    4440

TCGGCCTTCC TGCTTTAAAA AACACACGCG CAGAAACAAT GCCGTCCCGA TTTTCATCGG    4500

TATACAACAC ACCTTCGAAC TCCGCAAGAA ACCGCGCAAA ACGTGCGCGA ATCGGCCGAC    4560

TTGCAAGCAA ACGACTGAAC TCATGCAGCA GATCCGCATA CAAAACCACC ACGCGCTGCA    4620

AGGGAGTGGA TGAAGACACA GACGACGCAT GCACACCGAA GCTTGAAACA TCATCGGTCG    4680

GGTCAACCGC AGGAACTGCC AGAGGGGAAT TCAGCGTACA GAACATCTCC ATCGCGCGTT    4740

GTTCATCCAG CACTCCATAC TGCAATCCCA ACACAATTGA CTGGGCACGC GCGTACAGCT    4800

CCTGCTCAGG AGCaGAATCC CCTGCGCGCG TGCTTTGCTC ATATGAGGAT TGAGCAGAAC    4860

TCTCTCCAGG AGCATCTAAC GGACGCAAGG TAAAATACGT TTGCAGATAT CGAAATGCCA    4920

TGTTCGTTCG AGGTTCAAAT AGAGCCGCCT CTACAAGCAG CGATGGGTCC TGCTCCTGCC    4980

ATACAGAAAG ACGTGACAGG ATGGAATCTG CTATCTTTTT TGACCGCGAC GAGGGACGTC    5040

GTGAACGCTC TTGCGCAAAA AACAACTTTG CAAAGCGCGC GTCAAGcGCC CAACGTTCCA    5100

ACGCTTTTTC GATCAGCTCC TGCGCGTGCT CAACTTGTCC GAGCCCGTAA CGTGCCCGGG    5160

CAgCAACCAA TCTGCATCTG CAGACACCTG CTCAGCCGTT GCAAGAAGTT CCAGTGCACG    5220

GGCGTGCTGC AACGTATCCA CACACAGACG AGCATAAAAA AGCCGCACCT CTTCTATATC    5280

GTACACACAC CATTGCATAT CCTTTGCCAC GGCACGGGCC ATCCACTGTA ACGCACGTGC    5340

GCGCGGCTGC TGCAGCGCGT AAGAAGCCCG TGCAGCAATA AATAAAAAGT CTGCTATTTG    5400

CGGAgcaGAA GCGACTCCCT GCTCTGCCTG GGACAAAGCT TCCTGCCATC GCCCTTCCTG    5460

CAGATACCGA GCCGCAACAC CTGGATGATT ACGTTCTAAA TCCTGTGGAG GTGCAGGTTC    5520

AGACACGCAC GAGGCATCTT GGATGCCAGA AAATGCACAA AAGATACTCA CTGCACAGAG    5580

TCTTCCCATA CGTGGGCACA TTCCCTTACT CATGAGGAGT CCTCTCCGCG TAACGATTTT    5640

GGTAACCAAG TGctGCGCCA TAGAGCGCAC GCAACACACC ATCCTGCTCT ATCATCGGTA    5700

ACACCGTGCG GTCCGATAAA GGTACGTGCC ACTCTGAGAA CATCTTTCGG ATCCCCTTAT    5760

GACCACCGCG GATGGAGATG GTGTCTCCCG TGCGATGGGT TCTGATATAA AAGGGAAAAG    5820

AAAACGGACC TACACCCACG TGGTCCTGTG CGCAACAGAC AAACACGCCG GCAGGACGTA    5880

CTTCCACAAG AArgTTCCGC ACGCACAGGG GTAGGCACCA GGACGCGCCA CGTAGATTGC    5940

ACTCACTCCT TGCTTTTCAG AGGAAGGAGG TGATCCTGCA TCCTGTTTCT TTGTCTCACG    6000

TGCTGTGTCC GACGCATGTA TGCAGGAAAA AAGCACATAT GCACCGGCAC GCTCTAACTG    6060
```

CAGCCCTGAA ACGTGTATCC GACGCACACC ATCAAAACGC GCACACCGTT CAAGCGCCCC   6120

GCGTGGCACC CGGTGCGAAA CTCCCAAACG AACGCAAGCC TCCTGCAAAA GAAAGAAGCG   6180

CAATATAAAT TCAGCGGCGA GAAAGTCCGA CCGAGGCATC CGCAGACGCG TGCCCAACGC   6240

ACGTGGTACT GGTTCCCACG CATGCGAACA ACCTTCACGC CACCGCGTCA AGGCAGCAAC   6300

ACAAAAGCTG TGTTCTGCAC TAATCCCCGC AAACGTTTTG TCTAAGCCAG AGCGCCACCC   6360

TGCAAGCACT GCATCAAGTG CAGGGATAAG TTCATGACGG ATACGGTTAC GCACATATTT   6420

CCTGCACGTA TTTGATGCGT CTTCGCGCCA ACGCACACCA CGCGTCTGCA AGAAATCTTC   6480

AACACACGTG CGGCTCACCT TTAGCAGCGG ACGCACGTAC CGTCCACGCG CAcTCGTATA   6540

CCTTGCAACG CGGAcGcGgC CGCTCCCTGG AATAAGCGCA TGAGCAGTGT TTCGTACTGA   6600

TCATCACGGG TGTGCGCGGT TAGAACCACC TGTGCTCCGC AGCGAGCAGC CACGTGGTCA   6660

AAGACCTTAT AGCGCAGTGC ACGCGCCGcg TCCTGCACAC CGCGGCCACG AATTTTAGCA   6720

CACGCGTGCA CCGCACCGGc AGAAATCTGC TGCACGAAAC ACGGAAGGGG AGGAGAAAAA   6780

CGAGCACACA GCGCACGCAC AAAACGCGCA TCGAGCGCAC CTTCCTGAGC GCGCAGACTG   6840

TGATCAACCG TGACCGCGCA CGCACACACC CCAAAGTCAG GAGCGAGCTC GTGCGCCGCA   6900

TAAAGAAGCG CAAsmGAnTC GGCACCTCCT GAAACCGCCA CGAGCAAGCA AGAAGGCTTT   . 6960

CTCGGCACAA GGAAATGCCC AAAGctACGC GCCACGTGGA CGAGCAGCGG GTGAAGCTTC   7020

TGCCTAGACT CACTCACCTA TAAAGACGGG CACGCTGCAC AGTGTGCCGC ACCGCGcgCG   7080

TTACACCGCG CACCATCTAG CCGGTCCTCG CGCCAGCGGG TGAACCCGCT TCGGAAGCAG   7140

AAGAACTGAG TGCCACAATC ACCGCATCAG GATCGCTGAG AACCACCACC GACGCGGGCA   7200

GAGGAACATC ACGCACACGG CGCACGTCGC CGGCCCCGAG CCCACTGATA TCAAGCACAA   7260

CACGGTCGGG CAAGTTGCGC GGCAAAGACT CTACCTCGAT ATATGAGAGC CCCTTTTCCA   7320

AGCGAGCCCC ATAGCGCACT CCTTCAGGAG AACCACACAA CTGCAGCCGG ATTCGCATTC   7380

GCAACGGAAC ACTCTCTTCA AcTGCGTAGA AATCCACATG CTCCACACGG TCACTGACCA   7440

TGTTATGCTG ATAGTCCTTA ACAAAAACGC AAAAGACCTC GCCACCATCC AGTTCCAAAG   7500

ACAGAACAGT ACTCCTGGTT AAGGCACGAA ACAATCTATc GAAgnTTTGt GCGCAAGTTC   7560

aAGGGGAACG GACACGCCCC GATGGTCATA CATAACCGCA GrCAAACGCC CTTCCTTTCT   7620

GCCAgCACAG CGGCATACTT CCCCAACTGG ACGCGCCTTT TCCCCTTCAA ACGCCTTTCA   7680

TCCACAATCC AATCCTCCAT GCACAGAAAG CGAACACGCC GCAAACTGGG ACGGTAGGAT   7740

TCGAACtACG GAATGACGGT ACCAAAAACC GTTGGCTTAC CACTTGCCGA CGTCCCAAAG   7800

```
ATATCCTACC CACATAACCG GACACGCACA CACCAACACC ACCGCTTTGC AAGCAGCTTA    7860

TGCGCGCGCC GAAGCTCCTC CTCATCCCGA TACAGACCAA AAACCGCGCT CCCGCTTCCA    7920

CTCATCGCTG TAAAGCACGC ACCCGCACGG GCCAGATCCC AACGCGCAAG GGCGACTACA    7980

GGGTACCGAc GCTGTACAGG GGCATCTAAG CTATTAAAAA ACCGCCACCG CGCACAATCC    8040

TGTGCATAGT GCGCAGAAAG CGCGGTAGCC CCACGCAGAG AGTACTGCTC GCCGTCGGCA    8100

GCATGTACGC CGCACGCACG CAACCTGTcC AAATCCTcAT AGGcCTGTGC AGAACCGCTG    8160

TGCAATCCCG GcCAGACCAA AAGCCCCAGA TAGCCAGTCT TTGGAACAAG GGGAACGAGC    8220

TGCTcACCAC CACCTAGcAC gCACGCAGcC TGGGAAGCCA GGAAAAAAGG GACATCAcTG    8280

CCGACACTAT ACGCCAcTTC TCGTAGAAmC CGAGCAGAAa GGGTCGTCCC AAACAAgTAT    8340

CAaGGCCACA CAAAaGCGCG GCAGCATtCA GCAGACCCCC CACCAAGTCC AGACCtGCaG    8400

GGATACGCTT CACTACGCGC ACGCGCACAC CATCGTGAAC GCCAGTTACC TGACAAAACC    8460

GCGCATACGC ACGGGTCAGC GTGTTTTCTC GAGGCAGAGC CATATAAGGC GAACACACCT    8520

CACACCGGCC AGGGATATCC AGGCGCGAAA GAGACAAAGA ATCCGCAAGC GTAATGCGCT    8580

GCATTACACT CTCAATCGAG TGAAGACCAT CGGCCCGAgT GCACCAACCC ACAGATGCAT    8640

GTTCACCTTT GCGTGAGgCG CAAACTCAGC GACTGCACCC GCCATTCTAT GACAAGCGGA    8700

CACAGCGTGT CAATTCCCCC TTCTCTCTAC cTGCACCCAA AACACAAGAG AAAAAATACC    8760

TGTGCCTATT AGGCACAGTT GACAGCGTGT GCGCTCCCCT cTACGATCCA CCCCTAGCTT    8820

TCACCATACC ACAAGCAGAG GTCAGCCATA TGAACGAGAG AAACAAGTTA CTCGCACGCG    8880

CCCTGTATTC CTGCGTTCCA CACGTCCAAG GCTCGGACGA CTACGAGGAC GACTTTGAAG    8940

ACAGCGACTT CCAGGACGGG GATTTCGATG ATTTTGAAGA CGAGGATGGC TTTGACGATG    9000

ACGATGACTT TGAAGACGAC GATTTTGAAT ATGAAGATGA GGACAATGAC CTAGACTTTG    9060

ACGAATAGGA CGCACGCGCG GGTGTGGTTG TCGAGGCGAC ATGATCGCAT TCCTGTTGCC    9120

TGTGATGCGA GACTGCTAAG AAATCTTAAT AAAAAAGTTT TTGATAAAGC GTGCGCGTTC    9180

GTCTGCCTTT TTCCAGTATG GGCTGTGGGG GAAGCGTTCC AGTATTGTCT TGTATGCTTC    9240

GAGCGCGAGG CGTACGTTTC TCTGTGCGCC GTTGATCTCA TAGGCTTGTC CACGCAGGAA    9300

CCACGCTTCG TCCATTCGTT CGTGAGAAGG GAACTGCGCA AAGAAATCGC CGAGCGAAgn    9360

GgGGCATCTC GCGCGTTTCC CTGTGCACAA AACTGGCGCG CTTCTGCTAG GTGATCGCGT    9420

TTTTCTTGAC CCTCTTTGTG AGCAGATGCC GGGACATGCG CCTCAATAGG CGCAgCTGAG    9480

GGAGCAGAAG GTTGAGACGC TGGTGAAATC TTCCGCGGAG AGTACCGCTC CGACACACCC    9540
```

```
TGTGCTACTG CATCCGACGG CACAGGGTCA CGTCCTCCTA CGGTGGGTTC CCGATCTTTT      9600

TTTTCAtCAG GACGGGGCGT ACCGTGCTGA GCTTTCTCTG CAACAGCAGT ATCCGTCTGA      9660

TCCTGCCGGA CAGGCGCTCC AGTATGGGCA GCGGCGCGCT GAGAACCAGA AGTTCCACTC      9720

TCTTCTGCTC TGCGCTCGGT ACCCGTTCCC GCAGAGATAA CTCAGAAACG ACAGTATCAG      9780

GAGGAGACGA CACCGTACGC CGGTACTCAG GCGCACGCAC CACACGCGCG AGCCCTTCCC      9840

GCTTCGGTAC CACCTTGACC GCAAGTGCGT CGGAGACAAA ATCACCCCGA AACACATCAA      9900

AATAGGAGAA CGCTAAGACA AAATCACCCT CTCGCTCAGC ACTAAAGGTA AAAAGCGAAT      9960

GCGAnCTCCT CCAACTTGCG CTGGTGATAG CGCAAACCAG GCTGCGCAGT ATGCTCGCCC      10020

ACGTACACCC AACCTTCGcC CGGaTACAAA ACCTcAAGTT TTTGcCCCAC TGcAAGcTGT      10080

aCCGcGCGCG AAACGgGGCT ACCTtCATCC TtCAGGgCGG TTCTTcAGGc ACCATCGCGc      10140

GTGGAGAATC CTcTGcCGGC TCAGGCTCAG CCTGcAcCTC CGcCTCACGA GGAGGsTCTG      10200

ATGCAGGGGG CGGA                                                        10214
```

(2) INFORMATION FOR SEQ ID NO: 50:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 660 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: double
          (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

```
CTAATGAAGG CGATGTTTTC TTCTGAAAAG GACCCGTGGT ACCTACTCGG CGCGGGGGTT       60

GCGTGCGGTT TGGGAATTGC CGCTTCGGcG CTTTCTCAAG GGCGGGCTGC CGCAGCCGGC      120

GCCGATGCGC TTGCAGAAAC AGGTAAAGGA TTTAGCCAGT ATTTGACTAT CGTTGGTTTG      180

TGTGAGACGG TGGCGCTTCT GGTGATGGTT TTTGGTATTA TCAACTGCTA GATGTGGTGA      240

ACGTTGTGGT ATAGCGCTTC GACCATGCTT TTGATAGACG TAGGGAACTC GCACGTATTT      300

TCGGAATCCA AGGCGAGAAT GGTGGCCGTG TGTGCGTGCG TGAGTTGTTT CGCCTTGCGC      360

CTGACGCGCG TAAAACCCAA GATGAGTACT CGCTTCTCAT CCATGCGCTT TGCGAACGTG      420

CGGGGGTCGG CCGTGCTTCT CTCCGTGATG CGTTTATTtC CTCCGTCGTG CCTGTGTTGA      480

CAAAGACCAT TGCAGATGCG GTCGCTCAGA TTAGCGGcGT CCAGCCGtTG TCTTTGGCCC      540

GTGGGCGTAm GArCACTTGC CGGTGCGCAT ACCAGAGCCA gTGCGCGCGG AAAATTGGCAC     600

TGACTTGGTA gCCAAmGCGg TGGCGGCCTA TGTGCAnTTy CGTTCTGCTT GCGTGGGTAT     660
```

(2) INFORMATION FOR SEQ ID NO: 51:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8648 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

```
ATTTCCACAT TACTCAATAA AAAGACCCAG GATTTTAAAA AAAAATACCG CTACACCGCG          60

GATGTACTTC TTATAGATGA CATTCATTTT TTTGAAAACA AAGACGGATT ACAAGAAGAG         120

CTTTTCTATA CGTTCAACGA ACTTTTCGAG AAAAAAAAAC AAATTATCTT TACCTGCGAC         180

AGGCCTGTAC AAGAATTGAA AAATCTCTCT TCTCGCTTAc GCTCGAGGTG CTCCCGAGGG         240

CTTAGCACTG ATCTGAATAT GCCATGTTTT GAAACGCGCT GTGCTATCTT GATTAAAAAA         300

ATACAAAACT ATAACAGCAC CTATCCTCAC AAAGCCATCC ACATTTCAGA CGATGTTGTC         360

CGACTTGTTT CTGAAAACAT TTCTTCAAAT ATCAGGGATC TTGAGGGGGC ATTAACAAAA         420

ATTATCGCTT TCATTGAAGT GTCGGGATCC ATCACGATAG ATATCGTTCC CTCTCTCCTA         480

AAAGAGTTCT TCCTCTCTGC AAGGCCAAAA CACATCACAG TAGAAACTAT TCTTCATGTA         540

GTTGCAGATC ACTTTAACAT TTCGTATTCA GATCTAAAGG GTAAGAAACG CAATAAAAGC         600

GTTGTTTATC CTCGGCAAAT CGCTATGTTT CTCTCAAAGG AACTGACAGA GCTCTCCACT         660

ACTGAACTTG GTATCGAATT TGGTGGCAGA GATCATTCAA CCGTCATTTA CGGATGTCAA         720

AAAATAGAAG GAGAAATTCT CACTAATCCT TCGTTACAGG CAAATCTTGA TTTGCTGAAA         780

AGTAAAGTTC AAGATTCAAT CCGCTAGGGC GTAGACACTG AATTCGATGG GGATAAGTGG         840

TGGATAAAaG AATATAAATT AGTCATTACA CTTTACTCAC GAATATCCCC CTTTTTTTAG         900

AGAAAAAATA TACTTTCTTC ACAaGCTTGT GTGCGGTTTT TGTTTGGTAA TTCTCGAGAC         960

ATAaGCACTT ATCCAGATAT TCACAGTTAC TATTATGTGA TACGACTACA TTCTTTATAC        1020

TTATAAGATT AATAAGGAGG AAACTAACTG TGAAAATCCT ATGCGAGAAA GAAGCCTTTC        1080

TGAAGGAAAT AAGCACAGCA CAAGAGGTTA TTTCAAATAA AAAAAACACG TCTATTTTTT        1140

CGAACGTCCT ATTAGCTGCT CAAGGAGCCC TGCTTACCAT CAGAGCAACC GACACAAAAG        1200

TTACCTTTGA AACTAGCATT CCCGTCAATG TTCTCGCCGA AGaCaACGAC AGTTTTTTGC        1260

GACAAACTTG TGAATGTTGT TTCTGCCCTT CCAACAAAAG AAATCGAATT AACGTTATGT        1320

GAAGAACAAC TTGTCATTAC cCCTCCAAAC AAAAAGATAA GCTTTcAGCT CAGAACCCTC        1380
```

TCGCATGAGa GTTTTCCATG TTTCCCTCAA AATGAAGGAG GCGTCTCTCT TGCTGTGCCT   1440

ACCTCCGATC TTAGAAACAT GATTAACCAT ACCGTTTTTG CAGTTTCAGA AGACAGTACG   1500

CGCCATTTTA TCAATGGCGT ACACGTTGAT TTTCAGTATG GAAATATTAT TTGTGTTTCA   1560

ACAGATGGAA AGCGGCTCGC CTATATAGAA AAAAAGGGAG AATCCTCTCC CCAATCCTTT   1620

TCGGGTGTTA TTGTGCCAAC TAAGATCTTA GGCATAGTAA ACCGTAAGCT TACCCCTGAA   1680

GGATCAGTGA CGCTATGCAT TACGTCGCAG CACGTTTACT TTTTTTTCGG TGGATATAAG   1740

TTTTCTTCTG TGCTTATTGA GGGGCAATTT CCTAATTACA AAGAGTAAT CCCTGATCAT   1800

CAGGAGCGTT CTTTTTGTGT TGGACGTGTG GAGCTAATGG AGGCACTTAA ACGAGTCTCG   1860

TTGTTGGTAG AACAAAAATC TCACAGGATA TTTATTACCA TACAGCAGGG TTTGTTGACT   1920

TTAAGCTCAA AAGCTCACAC TCAAGAAAAT GAAATAGGTG ATGCTCAGGA AGAAATAGCC   1980

TGTGCTTATA CAGGAGAAAG TGAGGTCATA GCTCTTAACT ATCTATACCT TGAAGAACCG   2040

CTTAAGGTTT TTACTTCGAA GGAGGTTCAA GTGGAATTTA CCGATCCTGC AAAAGCACTC   2100

ACGCTTCGTG CTGTACCAAA CACGGACTGC TTTCACATCA TTATGCCTAT GCAAACGGAG   2160

TGATTCTTTG CCTTTTCTCA CAGTGACTGC AATAAATTTC AGAAATCTTG CACATCACAC   2220

GATTGATATA TCCTCTCCTG AGGTTTTTTT TGTGGGAAAT AACGGACAGG GAAAAACCAA   2280

TATACTTGAG GTTCTATATC TTGCTGCGTA CGGAAATTCG TTTCGAACAC GCACCGAAAG   2340

CGAACTGTAT GCAACTCACG CGCGTTCGAA TGAGTATCGG GTAAAAGTTA TGTACCGCGG   2400

GGAGTATACC CACACAGTGC AGATTTTCTC CAAAAATGGA AAAAAGCGCA TTGAGAAAAA   2460

CTTGAAAAAA ATAAGGACAA AAAAAGAACT TATCAGCAGT ATTCCCTGTA TTTTGTTTTT   2520

TCATAACGAT TTGGACTTCG TAGTTGGTAC GCCAGAACGC AGACGCTTCT TTTTGGATCA   2580

ATCCCTTTCG ATGTGTAATC CTCTGTATTT GGAATACTTG CAAAAATATC ACGCACTAAC   2640

AAAAACAAAG AACAGAGAGA TAAAAGAGAA ACGCGTTCAG TTACTCGATG CACTGGATAC   2700

GCAAATTGCA ACCGTGGGTT TTGATCTCGT GCAGTGGAGA ACTCAGCTTG TCCGTGACTT   2760

TAACGTGATT TTTACTAAGT ATTATGAGCG CCTTGGAGAC CTTGCGCAGG TGCGCATTGA   2820

GTATAAGCCT TCATGGTCTG ACTCCTCAGT TGAGGAGATC GTACATTCTC TTTACAAGAG   2880

ACGTAAGCAC GATCTTGCGA TGGGGATGAG TATGTCAGGT CCTCATAGAG ATAAGATTCA   2940

CTTTACTCGG TCGCAGGCGC TTTTCATTCC TCAGGCTTCT ACCGGACAGA GGCGGTTGGT   3000

TTCGTTGGTA CTGAGGATGT CGCAGGCTGT GTTCTACACA GGaGTAACGG GAAAACTGCC   3060

CGTACTCTTA ATGGATGATG TCTTGTTAGA GCTTGATCCT GAGAAGCGGG AAAGGTTCAT   3120

```
GATGAGTTTG CCTCCGTATG ATCAGCTGTT TTGTACATTT TTGCCAGGGG AAGCGTACAG    3180

GCGATACGGG CGTGAAAAAA CGCGGGTATA TTTTGTTTCT GAAGGGGCGT GTCATGAATA    3240

ATGGTGTGAA TAAGCTATCG GACTTACTCG TGTTGACCAC TGAATATATC CAAGCTTCCT    3300

ATGAAACGGA GGCGTTTGAT GCGCATCGAG AATGGGTGTG TATTGTGGGT AACCCCGTTG    3360

CGTTACACAG CACGCTGGTA GATATCAGAA ATGGGAAAGT TGTGGTCAAG GTGACTCATC    3420

CTGGTTGGGC ACAATACCTT TTGTTAAAGA AAGACGAAAT TGTACATGCC CTTCGTAGGC    3480

GATATCCGTC GTTGGGAGTG ACGGGTATGA GTACGTACGT AGATTCTACC TCACGTACCC    3540

CTTCTGCGAA GAAGGACATG CAGGGACTTT CGGTATCAGA AAAGCAGACT CGTCCTGTGC    3600

CTGAACTTGC CGAGGTATTT GAACAGCTCC GAACGCTTTT TCAGGTGAAA ACGGAAGAAC    3660

CGTCACATTA GTTTTGCGGA TGGGATTCGA CGGATCTGTT CAAAGTCCAT AGGACTGCGG    3720

TTTTTCTTGC GTGCAGCCTA TGCACGACTG TGTCTCTCCT TGAACGCAgT ATGGCTTTGC    3780

GTTAGAATGC CCGCCCTATG GAAGAAATTA GCACCCCAGA GGGTGGCGTT CTTGTGCCCA    3840

TTTCTATAGA GACAGAAGTC AAGCGTGCTT ACATAGACTA TTCTATGTCC GTCATAGTTT    3900

CTCGTGCGCT TCCGGATGTC CGCGACGGTT TAAAGCCTGT TCACAGACGT ATTCTCTACG    3960

CGATGGAGGA AAAAGGGcTA CGCTTTTCAG GACCTACACG GAAGTGTGCC AAGATAGTGG    4020

GGGACGTTTT GGGAAGCTTT CATCCTCATG GGGATGCGTC CGTCTATGAC GCGCTAGTGC    4080

GTCTTGGGCA AGATTTTTCC CTTCGTTATC CAGTCATTCA TCCTCAAGGA AATTTCGGGA    4140

CTATCGGGGG CGACCtCCGG CAGCGTATCG GTACACCGAA GCGAAGATGG CGCGTATTGC    4200

AGAATCTATG GTAGAGGACA TAAAAAAGGA AACGGTTTCC TTTGTTCCCA ATTTTGACGA    4260

TTCTGACGTA GAGCCCACGG TTCTTCCTGG AAGGTTTCCT TTTCTTCTTG CGAATGGGTC    4320

CAGTGGTATT GCAGTTGGTA TGACTACAAA CATGCCACCG CATAATTTGC GTGAGATAGC    4380

CGCAGCTATC TCTGCGTACA TCGAGAACCC AAATCTTTCG ATTCAGGAGT TATGCGATTG    4440

TATCAATGGT CCTGACTTTC CCACGGGAGG CATTATCTTT GGAAAGAACG GGATTAGGCA    4500

GTCTTACGAA ACAGGTCGAG GGAAAATTGT TGTCCGTGCT CGCTTTACCA TCGAGACGGA    4560

TTCAAAGGGT AGGGATACCA TTATTTTTAC AGAAGTTCCG TATCAAGTTA ATACTACCAT    4620

GCTTGTTATG CGTATTGGGG AACTTGCACG TGCGAAAGTG ATCGAAGGTA TTGCGAATGT    4680

AAACGACGAG ACTTCCGATC GTACAGGsTA CGCATAGTGG TAGAGCTCAA AAAGGgTACC    4740

CCCGCACAGG TAGTACTCAA TCACCTGTTT GCAAAGACTC CCCTGCAGTC CTCTTTTAAT    4800

GTGATTAATC TTGCTTTGGT AGAGGGAAGA CCTCGAATGC TCACGCTCAA GGACCTAGTG    4860
```

CGCTACTTTG TAGAACACCG GGTCGATGTA GTGACTCGGC GTGCGCATTT TGAATTACGT 4920

AAGGCTCAGG AGCGCATACA CTTGGTGCGT GCGCTGATAC GTGCCTTGGA TGCCATTGAT 4980

AAAATCATCA CGCTTATCCG TCATTCGCAG AACACAGAGC TTGCAAAACA GCGTTTGCGT 5040

GAACAATTTG ACTTTGACAA CGTGCAGGCG CAGGCGATCG TAGATATGCA GATGAAGCGC 5100

TTGACAGGTT TGGAAGTCGA GAGTTTGCGT ACGGAATTGA AAGATTTGAC GGAGCTGATT 5160

TCTTCTCTGG aGGAGTTACT TACTTCTCCC CAAAAGGTCT TGGGAGTTGT TAAGAAAGAG 5220

ACGCGTGATA TCGCAGATAT GTTTGGGGAT GATCGGCGTA CAGATATTGT GAGCAATGAA 5280

ATAGAATATC TGGATGTAGA AGATTTTATC CAGAAAGAGG AAATGGTTAT TCTTATTTCC 5340

CATCTTGGTT ACATTAAGCG CGTTCCAGTG TCTGCGTATA GAAATCAGAA TCGGGGAGGA 5400

AAgGGCTCAA GTTCAGCGAA TCTGGCGGCT CACGATTTTA TTAGCCAGAT ATTTACTGCA 5460

TCAACACATG ACTACGTGAT GTTTGTCACG AGCCGTGGGC GrGCCTATTG GCTAAAAGTA 5520

TACGGGATTC CTGAATCTGG TCGGGCGAAT CGTGGTTCGC ATATTAAGTC GCTTCTCATG 5580

GTAGCGACGG ACGAGGAGAT CACGGCCATC GTATCTTTGA GAGAGTTTAG TAATAAAAGT 5640

TATGTTTTTA TGGCTACTGC GCGAGGTGTA GTTAAAAAGG TAACTACTGA TAATTTTGTG 5700

AATGCGAAGA CGCGCGGTAT TATAGCGCTT AAGCTGAGCG GAGGTGACAC GCTGGTGAGC 5760

GCAtGTTGGT GCAGGACGAA GATGAAGTAA TGCTTATTAC GCGTCAGGGA AAAGCATTGC 5820

GCATGTCGGG GAGGGAGGTG CGCGAGATGG GTCGCAATTC CAGTGGGGTG ATTGGGATAA 5880

AATTGACGTC CGAGGACCTA GTGGCGGGGG TTTTGCGAGT AAGCGAACAA CGGAAAGTAC 5940

TGATAATGAC GGAGAATGGA TATGGTAAGC GGGTCAGTTT TTCAGAATTT TCTGTACATG 6000

GGCGAGGGAC TGCAGGACAG AAGATTTACA CACAAACGGA TAGAAAGGT GCTATAATAG 6060

GTGCTCTTGC TGTTCTCGAT ACAGATGAGT GTATGTGTAT TACTGGTCAG GGAAAAACGA 6120

TTCGCGTGGA CGTGTGTGCA ATCAGCGTGC TGGGGCGTGG TGCGCAGGGC GTGCGTGTGT 6180

TGGATATCGA GCCATCGGAT TTAGTAGTAG GACTTAGTTG TGTAATGCAG GGGTAATGGG 6240

CTCTGGGGTA TATTTCTCCG TGAGTGGCTG TGTATATGTT GTGAGTATTG TGGATAATGT 6300

GCGTGCAGAA GTTGATGTTT CACGTGAAAC TgTsGGGATG AGGAGTGGGA TCAAATCTAC 6360

CCTAATTCTG GAGGATTATT TGGGTTCACG TTCATGTAAA CTTTATGGGG GTTGTGTATG 6420

GGGACTCGTG TCAGATTTTC CTTCTGCGGT ATTGCAGGTG TATGTTTACT CGCACTAGGT 6480

TTTTTAGTTA GTTGTTCTTT GCAATCTTCA CGAAGCGCTA CAAAGAAATC TGAGGCGCGG 6540

AGGACTTCTT ATCGGATCGG TCTCATGACA AGTACGGGAT CTyAGTCTGT AGATGATGTC 6600

```
CTTGCGAAGA CACGCCTCGT CAGTATCTAC GGAGAGGCTC GTGGGGAAAC GGGTGGAAGG    6660

ATTGTCCATG TTACTTACTC CGATAACTTC TCCCACGACC ATGAAGCAAC CGTTTCTAAG    6720

TTGCTTGCAC TCGCTGAGGA TTCGACTATA AAGGCCATTG TGGTTAGTCA GGCAGTTCCC    6780

GGCGTTTCAA AGGCGTTTGG GATCATTAAG TCTAAACGTC CTGATGTTTT GCTTTTTGCG    6840

GGAGAACCAC TTGAGCCGGT AGAGATGCTG CAGGAGTCTG CAGACATCGT GGTCAGTCAG    6900

GACTACTTGT TCGGTGGATA TGCCGTTCCG TGGGTTGCGG AAAGGATGGG GGCGCGCACA    6960

TtGGTGCATG TCTCTTTTCC CCGGCATATG TCCTACCCCG GTTTGAGGGT TAGGCGTACG    7020

GTGATGAGGG CAGCATGTAC CGATTTGGGA CTTTCCTTCG CACACGAGGA AgCGCCTGAt    7080

CCTGTAGAcG GTGTCAGTGA CGGAGAACTT GAGGATTTTT TCCACAAGAC GATTGTGAAG    7140

TGGATCAAAA AATATGGCAA GGAAACCCTG TTCTAcTGCA CCAATGACGC TCACAACAGG    7200

CCGCTCATCA GTGCCTTGTT GAAATATGGC GGTATGCTAA TTGGTGCAAC CATCTTCGAT    7260

TACGCTGATG CGCTCGGGGT GCATTATGCT GAGCTTGAAG ACGTGTATAA AATACGAGAG    7320

AAGGTTGAGA AGTCATTGGk TTCTTCGGCG CAGAGGGGCG CTTTGGATTA AATTTAAATG    7380

CACAGGCATT TACGGTGACC ATGGGTTTTG TGGAGTATGC GCGCAAAATC ATAGATGGCG    7440

aACCGCGTAA AGATGATATG CGTGAAGCTC TTGCCGAATC CTTCGACTTG TTTACGCGTG    7500

ACGCACATTG GCGTATTGCT CCTTACCTAA GACTGAAAAC GCACGAAATT GTTCCGAATC    7560

ACGTGCTGGT GTATACGGAC ACATACGTCC TGGGTAAATT TACCTTGCCC GTCACAGACC    7620

AAGTACTCCC AGAAGGGTAT TGGGCATTGA CCGCTAAGGA ATAAGAACTC CGTTCGGGTT    7680

TTCTGTTTGT AGCCGGGGAG ATGGATCGCT TTCTCTGTTT GGCAATGTCG CCGTCTCCCT    7740

GGGTCACCAA GTGATCTTGC ACCCTAGAAA GAGTGAACCG GTGTATCCAG GCCAGCTCCA    7800

GTTCTCTTCT ATCAACATGT AGGGATCCTG TGAAAGCAAC CCTTGCTCCC ACCGCACGGA    7860

AAACTCCACA GGTTTGATAG GACTTGCACG CAGCTCAACA GCGTATTGGA AACAAAGTTC    7920

TCCCTTTAAA TTGCGCGTTC CTTTGAACCC ATTGAAATTG AATCGGTTGG TTGCCATATA    7980

TATGTGCGCA CGTGGTTCTA TCCACATACT ATCGTAGCAC GGTATGCGGT AGCCTACCCA    8040

TGCATTCCCC ATTATCGGAA GGGCTATTGA AgCTGCGCCT GTTGCTATCG CGTCTGCCGG    8100

TAACCCTGCC GCGCGTGCTA CAAAGTTTGT GGCACTATTC AAGAAATTAA GAAGACCTAA    8160

AATACCTACT CTGGCAGCAT TGGCGACCTG CTGTGCTACC CCTTGGGCAG GTACGACGTC    8220

TGGGGGGAGA CCTCCGTTGT CAAGATAACT TTTGTAGCCC AGGGGAAGGT ATACACGTGC    8280

TTCTATGCCT GCGTTTAGGC CGTGCAAGGC ATGGGTGTAA TCATCTCCCG AACGAGTTTC    8340
```

TAGTCTGAGA AACGCAGCAA AGTCCGTGTA TTGAAAAGTT GACTTTACAA AGGGACCACT   8400

CCCAAAAACA GACGCCGCCC CTGTTGCGCC GTATACACCT CCTGAAAGCC AACGcCACTg   8460

cGCTGTAACC AGCGCGTCTA TGCCTAATGA GTCCACGTGC TGTACCATCC AGCTGAATAT   8520

ACGACGCACC GTCCGTAATG ACGGATACGT CTTCTCCGCG AGTTCTAGTG CCTGTTCGAC   8580

GACACGTGCT CTCGcACTGT TCGTATCCCG GTAGGTATTT CCGGCATCCG nAGCTAAAAT   8640

GAAGCGGA   8648

(2) INFORMATION FOR SEQ ID NO: 52:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 6993 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: double
       (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

CACCAnCGTC CCGnATCCAG TTCCACGCAC ATTGGCAACG GCGCACAAgC GCTCTATCTG   60

ATCTTTGTGT ATATCAGAGA AAAcGCGAGC ACTGCAGAAA TATCTCCCGA ATTAATTTGC   120

AATATATTGC ATACATGCCG GAATGGTACC TGGTACGAAA TGCACGGCGG CATACCACGC   180

ACTTGTGACA ATTCGTAGAT CCTTTTATGC CGCATAAATT CGTGTTCGCT TTTTGCCGCA   240

TGTATTCCCC ACGCAACACG CTCACTTTTA TCGTAATCCT CGTATATCTT AAGAACATCA   300

AGATCAAAAC TAATGGAAAA CTCGGTATTT GGACGTGTGG AAACAAAAAG GTATCGCAAT   360

ACTTCAGGCT GATATACTTC AAGCACATCA CGCAGCCCAA CCACTTTTCC CGCGGACGAA   420

GACATCTTCC CAGGCAAACC TTTTAATCCA ATAAAATCAT AACGAAAAGA AACAGGCGCA   480

GGCCAGTGAT AAATGTGATC AGAAATTAAA CGCGCAGTGT CAAAAGAACC TCCCTGAGAA   540

TGATGATCCT TCCCTGCAGG CTCAAATACC ACATGCTCCT TACTCCACCG CATAGCCCAA   600

TCAACGCGCC AGcTAAGTTT TACCGCAGAC GTCTGGCGTA AATCCACCTG CTCCCCATGC   660

CCACACTCGC AATGATACTG AAGACACCAG TGGCTATCCC ACGCATCAAC CGTGGTGCAG   720

TCTTTATGGC ACGCTGTACA AAACACCGAT ACGGGCCAAT ACGTTCCACT GATTTTATGC   780

TGCTCATCTC GATATTCGTT TAAAATCGCT TGAATACGGT GCCGATTGTC GAGCGCAATC   840

TTTATTTCCT GTGCGTATAC CCCCGCCTGG TATTGCTTTG ACTGATAAAC GTATTCAGGA   900

TAAATACCTA CCTCCGGGAG CGCCGATTCA ATTTCCCGCT cATGGTGCCg CGCGTAcTAT   960

CTTCCTGCTG AAAGGGATCA GGAACTGAAG TGATAGGCAT GCGAATATAC TGCTTCAATT   1020

```
CATCTTGAGC AGGTACATTG TCGGGAATCC TACGAAAAAC GTCATAATCG TCCCACGAAT    1080

GTACAAAGCG CACTGATTTC CCCTGGTCAC GCAGGcCCGC ACTACAAGGT CAACGGAAAT    1140

AATCTCTCTG AAATTACCAA TATGTACCGT TCCTGAGGGG GTAATCCCCG ATGCACAGGT    1200

GTATTGATCA CAGTCAGCAC GTTCCTgATA ATCTgTGCGC AACCTgTCAG CCCAATGAAG    1260

TGACTTTTCA CAGATACTCA TGATCCTTCC TTACAGGTAC GCAAAATATT TTTAAAGCCA    1320

AGCCAGTACG CACTTCCCAC GCGCACATCC TACACTTCCA CATAGGTGGC AGTGCCGAAT    1380

ACACACAACA GCCATACGGT GAGTGCTCGC ATCTGCCTGG CAGACCGTAT GGTCACTCCC    1440

GCGATACAGG GAAAATAAGA ACTGCGTGCG TGTCAATACC GCGCGCACAG CATGAAACCT    1500

ATTGCTAGAG GGTGCATCTC CTTCTCAGAT TCCTATTCAT GCCGTATCAC TTGATGCGGC    1560

TCAGTATCCG GTCGAACCGC CACCACATAA CGACCGCCTC TGTGTCCGTC GCCAGCACTG    1620

AAGAAAGAAT CCCTGTGGGA GAAAAATGAA ACACATTGTA CACCAACGTA TCACTGCGCA    1680

TATGTATGCT CCTTTGACGT ATCGTGAATG AACGCAAGTC TAACAATAGC GCGGCATATC    1740

CGAAGCGGTC CGGACTCACA CACAAATAC CTTGtGCGCA CTCACCCCGA GCAATGAAAA    1800

aGCTTTTTTA TAaGGTCCcC TTTCTTGGTT TTCTGTACCA GACACCTCGT GCGCGGGAAG    1860

ATCCACTTTC CACTCGTACA CCCCCGTTTC AAGAGAGATA AAATACACAC TGCTTTTCGC    1920

ATAGCGAACA CTACTATTTG CTCCCGTAGC TGGATCATGC ACTGCGGTAT AGTAATCTAC    1980

CTTTGCAATC AGCCTCCGTT CACTCACATC AGGCAGTACC CGCTCTACGA ATGCATACCC    2040

TTTCTCTTGC GCAGcAAATG ACGTCGGAAG CGGAGAAAAC GGCAACGCAC GTTGGTGTAT    2100

GGGTACTCGT TGTGCATTGA ACCAATACAC TCTCATTGCA TCCACACTCC TACACACCAC    2160

CACGAGCTCA TCTGCACTAT TTACATACAC ACCTTCAATA GCAGGAAAGG GTGTGCCCCC    2220

TATCCCCTGC TGTCCAATAG CATGCATGAA ACGTCCTTCC TCATCGAACA GCAATATGGT    2280

ATTCACCAGC GCAAGGTTTT CTTCCGGATC ATGCTGCACG TGTTCTGGTA ACACGGCATC    2340

TAtACGTACA GTGTATTCTG CGAATCTACG GCAAGAAACG TTGGCGCATG CAGCGGATAG    2400

GGCACGGCAC GACGCGTAgT AATTGCTGCC GATTGCAACC CTTCAGAGAA CTGAGGAGTC    2460

ATTGCGTTTT TCTCGGGATT AAAAATAACT GCAAGCACAT CCCCAAACGA AtCATTCGCA    2520

TGACcTTGCC AGTTGCAGCA TGTGCCACGT AGAAAATACC GTCTTTCATA CACAGCTGTA    2580

TATTAGACTG CGATTGCGCA TACCCCGCGT CGGGGAAATG CAGTTGATTT TCAGCATCCC    2640

CGTACGTTAG GGCGAACAGA CGTTGCCCAT GCAATTCACG CCCCATCCAC CGTGTGCAGG    2700

AAGACACCAA CACGAGAAAA CTACCCAGCA AGAAAAAAGT AAAAAATCCC AACCGCAACG    2760
```

GGTGACGCGT CACAACGCTA CAGGACACGA GGATAGAGAT ACTCAAAACC ACAAAACGGC   2820

ACCAAAGCTT GCGGnAATGC GCACGCGACC TTCCGCATCC TGCCCATTTT CTAGCAGCGC   2880

AATTAACACG CGAGAAATTG CAAGCGCCGT CCCATTCAAC ATGTGTACAT AATGCTTCTT   2940

CCCTTCTGCA TCCTTATAGC GGACATTTAA GCGCCGCGCC TGATAGTCTG TGCAATTCGA   3000

CGCAGAAGTC ACCTCTCCCC ACGAACCACC CTGGCGTCCA GGCATCCACG CCTCCAAATC   3060

CCACTTGCGA TACGCAGGCG CACCCAAATC TCCCGCACAC ACTTCCACCA CACGAAAAGG   3120

AATTTCCAAT GCAGTAAAAA TCTCTTCCTC AAGCGACCGC AGgCGTTCGT GCAGGCACTC   3180

AGAATCACTC GGTGTACAGT ACGCAAACAT TTCAAGTTTG GTAAATTGGT GCACGCGATA   3240

AAGACCGCGA GAAAACTGGC CTGcAGCACC AGCCTCTTTa CGAAAACAAT GCGAGAGCCC   3300

TGCGTATAAA CGCGGTAAAC TCCGCTCTTC AAGAACCTCG CCTGcATGGT ATGCCCCCAG   3360

CGTAATTTCT GCAGTTGCTA CTAAACAGCG GTGTTCTCCC TCAATACGAT AGATATTCGA   3420

TCCACTCCCC CGCGGATTAA AACCCAAACC ACACACCATA CCCTCACGAG CAATGTcAGG   3480

AGTGAGAAAT GGCACAAAAC CGCGCTCTTG TAAAAACTGC AAACCAAACA TAATCAATGC   3540

CTGTTCAAGC AGCACCCCTT CACGCTTCAG ATAATAAAAC TTTATCCCCG AGACCTTTTT   3600

CCCCGCTTCA AAATCAACTA TATCCAGCAA GCGCGCTAAT TCCACGTGAT CACGTGGcGA   3660

AAAACTAAAG CATGGAGGCA CCCCACAGCG CTTGATTTCG AGATTATCAC TGTCTGATCG   3720

ACCATGGGGA GTGCACATAT GCGTCATGTT TGGCAACGCT TGCGTTGCAG ACAAAAGCTG   3780

ATCGGAAATC TGTACCAATA GACGCTCGCT GTGAGCAATG CGATCTTTTA GTGCTCTGCC   3840

CGTTTCAACA CACGCCGAAC GCGCAAGcGC ATCCAAAGAG CTTTTCATCG TCTGTGCGTT   3900

CTCATTACGC GCACGTTGTA ATTCTTGCAA CTCTGCTAAA AGCTTTACGC GCTGATCATA   3960

TAAGTGCACA ATCGCGTCCA CATCTGCATG CACGTTCCTG ACCTTCACAT TTTCTTTTAC   4020

TGCATCCACG TTCTCTCTAA TAAACCGATA ATCAAGCACG CGCCTTTCTC CCCTTACTTA   4080

TTCTGAATGT ACAAGAAAAA CGACACTCTC ATCGAaTGCT GCGCAGAAGC GCTAACAACA   4140

TACCCATCGC CCCATCGTGT ACGAATGTGT CAGACGTGGT AGCCGAGCTG TCCGGAAGGC   4200

GCGGTTCACA TATTACCGCA TCTCCCACGC TGATGTGATC GTAGTATCCT ACGCGCCTGA   4260

GCACGCCTTC ACTCATATCC TCTGAAACCT TTGTCACGGT AAAATGCCCC AACACGTCCT   4320

CACGTCGATA TGCAAGCCCA ACACCTTCTT TCAGCACTGA AAcTGCGCGT CTTTTTACCA   4380

CTTCGAGAGA CTTACCTTGT AACTCTGCAT CTTGTGTTCC AAGATCAATC ACCGCcTCCG   4440

ACTGGTGACG GCGCACGACA GTACCCATAA TTGGCAAACG ATCGTTGAGC ATCTGCATGA   4500

```
TCCTACGCAA CACACTCTGA TACCGATCAT TTCCCGAGCG ATACGCATCA AAGGTGTGTG    4560

CCCGAGATCC CGTCGATGCA ACATACAATT CCAAACGCAC GCGTAAATCC TGACCGTGCT    4620

CCTGCATTGT GATGAGAGCA AAATAATCAT CACCAGCCTC ACGCGCCGTG CGAAAAGCTT    4680

CTCTATACGA GTGAGAACGC GCGCTGTACC CAGTTACTTT TTAaCTGCGG TTATAGGCAA    4740

ACGAATCTTG CACTGCGTCA GAAAGAAACG CTCAGCCTCA GGATGCAATG CATTCGCAGG    4800

ATCAGGATGG TAAAAAGAG ATATTGAAAG ATGCGCCTTA TCCAGATACA AGGCATCAAC    4860

ACGCCACCTA TTTTTAATTG AACGTGCATG CGTCTTCTCG TATGCCTCTA CTGCATCATT    4920

GATGCGTGCA CTGCTCTTTC CAATAGACTG TAAAAACTTT AATTGCTCAA GGGAGCGCTC    4980

AGGATATCCA AAACGCAGGA GCAAACGTGC GTAAgCTTcA CGAGCAGCAC CGTCGTACGG    5040

GTACACCTTT AGTGCGCGCC GATACTCATC CAGAGCCTGC CGACTCATAT TCCGACGCGC    5100

GAAACCGTCT GCCTTTTGCG TGTGAAAACG CGCAAGTTGC ATGCGATACT CATCTTCGTA    5160

TTCAAGGTGA ACAATCGCGA TCTCTTCTAG CAAGATGCGC ATTAGATCGT CACGTGGATC    5220

TACTGTCAAC CCAACTTTTG CAGTTGCAAG CGCCTCAGTA TGTTTACCCA ACTTCAGAAG    5280

GGACAGTGTC TTTACATACC AGGCATCCAC TTGCGTTCGA TCCGCCCTTA TGCGTTGATC    5340

ACACTGAGCC ACCGCGCGCT CATACGCGCC GCGCGCATAT AAAACTGCTG AAAGAAGCGC    5400

ACGGGCACGT GGATAAGCCG ACTTAATGTG GAGCGCCCGC TCCAAATAAC GCTCTGCATC    5460

TTCATAGTGT GCCCGAAGCG TTGCAAGATA CGCGGCAAAA AAATGCACCT GTGCATTATC    5520

ACCGTGATAT TGCAACGCAC GTTCAACGTA CGTGAGCGCA CGCGGATAAT GACCAGCCTC    5580

GTACGAGATA AGCGCAAGcg ACAACAACGC CTTGCGATTC TCTGCCTGAC GCTCCAGCGC    5640

TGCTTGGTAT AACAGACGCG CAGAGCTCAG CCGTCCCTTT GACACCTCAA TCTCTGCCAA    5700

ACCAAAGCGA GCATCTACGT CATTCGGATA GCGCGCAAGA ATTTCCTCAA AAAGACTACG    5760

CGCCTGATCC AACTCACCTT GACCAACTAA ACTGAACGCG CACAGCTTTT CAAGGGAAAG    5820

ATCCTGCGCC ATGAGTTTTT GCGCTTTGCr CACATGGTGC AACGCCTGAT CATATTCACC    5880

AAGTGCGTAG AAACACTCGG CAAGACCACG ATATGCAAGG TTGTAAGAAG CATTTTTTTT    5940

TAATGCTTCT TGGTAGAATT CGATAGCAGC ATGCCAATCC TCCTGcACAT GGGCCTTTCT    6000

TCCTGCTTCG TAAAGCTGCA CGCCCGTCTG AGCAAACACT ATGCTGCAAA GCGCACCGTA    6060

ATACACGCAC AGCAGGCCTT TCATGCTTTT TCCCTTTCCT CTATGGCCGC TCAGATACAC    6120

ATGCAGGCTC GGAATCACCC CGCAGaCAGA TACAATCTTT ATAGTATTTA TCTTATGTGC    6180

TTCCATGTCT TGGATAATAA AATCGAAGGC ACCCCACGAA ACCTTCTCGT ACTTTACGGG    6240
```

AATTTTTCCA AAAAGATTAA ATACAAATCC ACCTAACGTT CCAAACTCTT GAGAAGGAAA   6300

AACAGTATGC AAACACTCAG ACAAATCTTC CAAATCCACA CGCGCATCGC ACAACCACAC   6360

GCCCTGTCCG AGCGGTTCGA TATCCTCCCG CTCGTGGTCA AACTCATCCT GGATATCCCC   6420

AACAATCTCT TCAATAATGT CTTCCATGCA CGCAATACCC GAAACGCCGC CGTACTCGTC   6480

CACCGCGATC GCAATGTGCA CGTGCCTGCG CTTAAACTCT CGCAGAAGAC TGTCAATTCG   6540

TTTGGACTCG GGGACAAAGA AgGsTTACGC AGCAGTCTTT CTAACCGCAC CTCCTGTGGC   6600

CTTCCAAACA GCTTTATTAA ATCTTTGACG TACAGCACAC CCACCACATT ATCAATAGTT   6660

TGTTCGTAGA CAGGAAAGCG TGAGTGTCCA CTCTCGGTTA CCTTTTCAAC GAGTGTTtCA   6720

CCGCTCATAG AAAGCTCAAG AAAATCCACG TCAATACGCG GTATCATCAC CTCGCGCACC   6780

GAAGTGTCAG AAAGATCCAC TATAmCGCGG rTCATAtCCT GcTTTTCTTC ATTCAGCGGT   6840

TGCTGAAAAA TATGGGTAAC AGCGTGCCTG CGCCTCAACC AGTCTATGAC TCCCATGGTA   6900

TACCCGATGA TAGCACCCGA CACGTGTGCG CCAGTATGCG CTCCTGCAAA CGCAACATCT   6960

CTTGTCCAGG GnTCCTnCGA TCAGACTCTA TAA   6993

(2) INFORMATION FOR SEQ ID NO: 53:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 5460 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

TCGCGnnAGT CAAAAACGGC AACACTGAGT TTTTGCTCAT TGGGGGCAGC CAGGGGTACA   60

AGGAAATAAA ACTGGAAACG GGGAGCGGCA GCGGTACCGG CTGCCTGAAG GCAGAGAACG   120

TGCGCGGTCC GGAACAGTGG GGTGAAGACA GTGTCACTCC CAAGGATAGG GTAAGCCAAT   180

ATGAAGGCAC CATCGGCCGT TTCGCAATCA GCGACATTTA CACCGTTGAG TCCACGAGTG   240

GAGCTGGTGG CACCAACGGC GGCACTAATA AGCCGGACGT GTAtGTGGTG GTGGGGGATT   300

CACAAGACGG GTATACGGGC CTGTGGAGAT TTGACGCCCA GAAAAAGGAG TGGAATCGGG   360

AGTAGCCCGG GCGGATGCGT GCTGCAGGGA GGCGCGGGGC GGGAGGCCGC GCGCCGGTCA   420

TCTTTACGCT TTGATAAAAA ACAGTTCGTG AATGGCGCGC CCCTGCGTCT GCGCCTTGCG   480

TTCAAATTCC GTGGCGGGGC GCCAGGGGCg cGCACCCTGC GGTGCCCACG TGAGCGAGGG   540

CGTGCGCGCA AgcTCTTCCT GCGCGCGCCG TGCGTACTCG GCCCAGTCGG TGACCGCGTA   600

```
TAGGTAGCCA CCCGGTGCAA GGGCGCGCGC GAGTAGGTCT GTGCGCGGGC GATACAGCAG    660

GCGCCGCTTG TGGTGCCGCG TTTTTGGCCA CGGGTCTGGA AAGAAAATGT GCAGGCCTGC    720

AAGTGTCTGC GGTGCGATCA TGGTGCGCAg caCGTcGAGT GCATCGTGCT CGATGATGCG    780

CAGGTTGTGT AAACGTTCGG CTTCAATTTT TCTCAGCAGT CGTCCGATTC CTGCGCGGTA    840

CACCTCGATG CCGAGGTAGG AAAGGTGCGG GTTGCGTGCC GCGATTGCCG CAGTTGCGCT    900

CCCCATACCA AAGCCAATTT CTACTACCAG CGGTGCAGGC GCGCACGCCG GAGCGACTGC    960

GTCCGTTTTC CCCTGCGGAc GGGaAAAGCA CCGGCAGGCG CAGAAGGTGC CGCCGGTGAA   1020

CAGAATACGG CAGCGTAGTC GAACACCGTG TTCTGATACG GGATnATCCA GCGGGACGCA   1080

AGGTGCTGGT AGTCGCGTTT TTGGCATGCG GTCATGCGGT TTGATCTGCG CGTAAAGGTG   1140

AGAACTTTCC GCATGCGTGC ACTGTCGTTT GTCATGGTGG CGCTTGCTCA GACAGGGCGT   1200

CTTCAGGATA TAAACGGTGA GGTTGTGAAA TAAAGCGCCA GGAGCGCTGA AAGTCCTCAA   1260

CCACGCACTG CAGGTAATAT GCGTCCTGGT GTGTCGTTTT CAGTGCGTCC ATGGACGCCT   1320

CGAAGGTGTG TACTCGGGGA GAGGAAAGGC GCGCATTGGG CAAAGAAGTT AACGAAAGGG   1380

GATGCTGTGC CACAGTGCGG TAGTCGCGCG TCACAGCACC AGGGAAGTGT GCGTGTGCGC   1440

AGTAAAGGGA CCGACCAGTG CCGGCTGCGC AGGCGCCGAT AGCGTCCAGC GTTGCGCGTC   1500

TACAGAAAAG GAAAACGGAA AGGGTTCGTG TGCTGGGTAG CCTGCGCCTA GGGTAACGCG   1560

CTGCTGTGCA TATGTGCCTG TGCTGTCGCG CACGTACACC ACGTACGTTC CGGCAGGAAA   1620

ATCTCCCCAC GGATACTGCA TCCCCCCTAC GCGCAnATtA CGGCGTCCAG GATGGAGTTC   1680

GTCAGTGGCA ATGTACACGC GCGTGTCTTT TTCCCCAAAG ACCCATCGCA TTCCGGTGCG   1740

CAnTCCTcTA CTTCAAGGTA AAAGTACTCC TCAGGAGGAG CAGGGTACTC AAGCGTCACG   1800

TGCAGTGCCa GCGTTGCGTC TGTCTTTCCT GCTTTTTCGA AATACACTCG TTTTAGGGTT   1860

AGCCCCTCTG TGCGTGCATA AAAGGGCATA CAGGAAGAAA GTACTGCGCC CGCTACACAC   1920

GGCACCACAC AAGAAACCAG TACAAACGCG CACTGGCGAG CGAGAACCAT GCGATTAAAA   1980

CTCAAAAGAC AAATCCACGG TGGTGAACTC AGAGCGTGTA AAACGGTTAA GACTTGACTC   2040

AAAGCGCACG TTTGCCTCTT TGwCnCTTCG CTCAAATACG AAATGTAGTA CACGCCCAAG   2100

TTGGCGTGCT TGGCGGTATC TATAGTTTGC TGGCAGTAGT CAACAAGCGT GTCAGTGTCC   2160

GCATTTACCA CTTTTGCATC CTGTAAGCAA TTGCGCATTT CTTGTGGGTA ATCATTCCTC   2220

CCATACACGG TAATTTTTAG CTTACGGCGG GTACsGGCAG TAGACACGCT CCTCCCCTGA   2280

AGTTTCCACG AAATAAAGAG AGACTCAGCG TTGTGTTTGA GTTCTGCAAC AATACGCCTG   2340
```

CGGATACCTT CGTCGCAGAG CGCCTCTTGC GCGTCTTCAA GGTCTGGAAA GAGCTTTTGC   2400

ACCTCTTTTT TTAAATTCGT ATTCTCCTCA TTGATGAGGA GCTCTACAAG GACTAATGCA   2460

ATATACTCAG CAATGCGCGC CTTATTCAAA TATTGCGCAA GGCTGGTGTG AATGCTACTG   2520

AGCACAGACC CCATGTCCGG TGCGTGCTGA AACTTACTGA TGATAAACCA CGTAAAGGGG   2580

CGCAGGGTAT CCAAAAACTT TTCGCCGAGG AAAAGAAGCG TATTTTTCTC CTTTGGAGTA   2640

CGTTCTTCGC ACGCAGCGAG CGAAGCGTGA AGCAAACGAA GAATATGATG CTTAATCTGC   2700

GTGACGTGTG CTTCATGCTG TTTTAAATGA TTATAGATGA AAGGAGCATT GAGATTTGCG   2760

TGTTCGTCAA TGCTGCTCCC CGGGTTACTG CGATTCCACT TTTTAATTAC TTCAGAGTTC   2820

AGAATCTGTC TGAACACGTA GCAATCGTAT TGACGGTAGA GGATGGAATG CACGATGAGC   2880

TTTGAAAGAT CAATAATTTC TTGACGTGAG GAAGCAAACT CAGGCGGTGA AACCTCAATG   2940

AGGGAAACGT ATCCAGAGAG CAAAAGACCT TGCACCGTTT TAGGTGCAAA GGCATCGAGT   3000

GCGATACCAT AATCCTCGAC ATGCTCAGCG AGTTTTAATT TCATGAGCTT CCTGTTTTGT   3060

TTTATAAAAA ACTCGCTGCC CTCTTGCGTG AGTACGAGCT TGAGTGGGAG GTTCAGGATG   3120

CGTCGTTTAT GTTTTGAACT CATACTTTTC TGCGCTCCCG TTGAAGTGTG TCCGCCGTCC   3180

TTTCATTCTA TCAAGGAATG AGGGGTGGGG GATAAAGAGA TTCTACGTAG CAGACGAACC   3240

GCACCGATCC TTCCTGTGCA TGCAGGAAGG AGTCAGGAGG GGCGGTGGGG AATCAGAAAT   3300

AACCCAGAGA AAGGCTGATG TTCTGCGCAA AAGCATCCGG AGTGCTGACC GCTACAAAGA   3360

GGATAGGAGC TGCCGCACGT GAGGCGCCTG TGCATAATCG GGATTGCGAT CCAAAAGCTG   3420

CGTTAGTACC TCTTGTGCAT ACTCTTTTTC GTTCACTTTG ATTGCTAGTT TCCCAGCCTC   3480

TAGGTACACG TCCCACGCGC GTGCGTCCTG AGCGAGCACG GCACGGTACG CGCGGAGTGC   3540

TTCATGCGGC TGTCCAGCCG CTnACATACA CGCGTGCCAC TTGGCACTGT GCCTCGCGGT   3600

CTTGAGGCTG TGCGGTGGCG GCGCGGCGGT AGTGCGCTAT TGCCGTCTCC CACTGCGCGC   3660

GCAGGACATA CAACTTCCCT AGATTGCTGT TTACCTCAAA GTTTTGCGCA TCGTGGGCAA   3720

GAGCCGCCTG CAGGTGTGTT TCCGCCTCTT GCAATGCTCC CTTGTCCAAA TACAGTTTGC   3780

CCAAATTGTT GTGAGCCTTC ACGTGTGCAG GGTCCCGTGC TGCAGCCAGC TGATACTGTG   3840

TTAAGGCAAG ATCCACACGC CCTGTTTTTT kCGCTGCAAc aTnAcGCGTA GAGGAAGCGT   3900

GCCTGCGATC CATTCGCGTA GACGGCCtGC TGCGCGTGCT TTAACGCTTC TTCGTTCCTA   3960

TCGAGATCAA GCAGCACGGT TGCGAGGTTG TACAAAGTAA GCGCGTCCTC AGAATCCCGC   4020

TCTAGGATTT CTTGAAACAG ACGCACAGCC TCTTCCTTCG CACCACTTTT TGCCAGGACA   4080

```
CGTGyGCGTT CACGCGnGGC TGTCAGGTAT AGATCATGCG cTTCCGCAGC AGCCGCGCAT    4140

TCTTGCGCCT TTTGATAGTC GGCAAGCGCA gCAGACACTT CTCCTTGTGC ATCGTGCACA    4200

CGTCCAAGCT CTATcCACGC ACGCGTGTGC GTTGGGTTCA ATCGGATGAC CGCCCTAAAC    4260

GCCTGCAGGG CTAGATCGTG TTTTGCACGT ACCCTACAGG TGAGCCCCAG ATTAAAAAAT    4320

GCAGGCTCGA ACTTTGGATT CGCAACCGTC GCCGCATTGA ACGCTTCCTG CGCCTCAGCA    4380

AACCTGCGCA CTGCGAAAAG ACGCTTACCC AACTCATAGC TGTAGCGATA TTCGCGCCGG    4440

TCAAGCGCCG CGGCGcGTTC AAGCAACGTG AGAGCCGTTG TTTGGTCGTC ATCAGGctGC    4500

GCATCTGCAA TGCACGCGGC AAGGTAGTGT GCAGCTGCGG AgCGTGGGTT GAGCCTCAAG    4560

GCTTCCTTCA CATACACGGT TGCCGTTTCA AGTGCCCGTG TCCGCTCAAA ACCGTCACGG    4620

TTATCGTGCT GTGAAAGCGC GTACATAGCT TCTCCCATAC GGGTGTATGC GTCTGCTGCG    4680

AACACCGCGT CCCCGGCAGG GAGTGCACGT ATTGCTTTGT TAAACACACG CACCGCTCCT    4740

GGATAATCAC GTCGTTCCGT CAGTTCTTTT CCTTCGGAAA GCAACGCGTg CACGTGGTGC    4800

TGTGGCGTGG CAGTTTCTGC AGGTGCGCGC ACCTCCGGGC GCGTTGCAAT CTGCACGGCC    4860

CGCTTTATCG ATTTTTCAGG AGAAAAAGAC GCCGTGCGAG ACACTCCCCG GGGCGGGGTG    4920

AGAACACGCG CGCCCTTTTG CCGCGTGTCT TGTTCTTGCA TACGCTCTCT CGGCGTAGGC    4980

GCGTAAGGAA CTGGGCGCTG GGGAGCTAAC TCCTCTGAAA GGGTCTGTAG GAACTGTTCT    5040

TCATCACTAT TCCCTGAGAG CTGGACGTGA TGGTCCACAA GCACGCCGGG TTCCTCCCCT    5100

TGCTCTTGCA GCAGCTGGTT CGTCTCCAGC CAGGCGAGTT CTTGCTCAGA AACGCCCTCC    5160

TCCTCAAGGA GAGTCTCGCC CGCTGCACGA GGTAGGCGTG CGCGCACCAC CCCCCGGGAA    5220

AAGAGACTGA ACCCGGCAAC TACCGCAAGA AGCAGCACCA GCCCCGCGGC GAGTGCAATG    5280

AACGTCTTGT GCACATTATT CAAGGTTGTG TTCCTCCTGA TAGGGACGG TGTCCTCCGA     5340

TCCAGTGGAG AGGGTAnGCG CGTCCTCCGC TTGTTTCAGT CTAAGCGCGC GCTTGAGAGC    5400

TTCAAACTCC GCCTCCTTGC GCCGGGnTTC CTCCGGCTTC CTTGCGGCGG GnTTTCTCCG    5460
```

(2) INFORMATION FOR SEQ ID NO: 54:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10461 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
AAATCGTGCT GATACTGCAA CTTAGTAGCG ACAGGAGTAA TGAACCAATT TTGCACCAGC    60

GAACCATAGA AAGTAGTGAT AAGCGCATTG CCATGTTAGA TCCCAGCGAG GACTTGTCTT   120

CAAGCGTTGC AAGCATACCG ATAAGCCCCA TAACGGTGCC CAGCATACCA TATCCGGGCG   180

CGAGcGCAgC CCAGGAGTTC AAAAGGGAAA TCCACGTATT GTGCCGATCC TCCATGTGCG   240

TCAACTCGCT TTCCATCAGT GCCTTGATCG CATCTCCGTC CACACCGTCT ACCACGTTCC   300

GCAAACCAGT GCGCACGAAn TCATCGTCAA AGTCCTGAAT TTCTTCTTCG AGCGCAAGTA   360

AACCGGTGCG CCGACTTTTC TCAGCAAGCG CGTAGAGCCG CTGGACAATC TCCCGTTCGT   420

GAAAATCCGC CGCATGAAAA ACGCGCGCAA TTACCCGAAA AACACCCACG GCATACGAAA   480

GCGGATAGGT GAGAAAAAGC GTTAAGTACG AGCCCCCCAC GGTGATCAAC AATGACGGTA   540

CGTGAAAGAG CCCCCTCGCA GAACCACCGA GCACCGCACC AAAGATAATG ATGGCAAAAC   600

CGCCGAAAAG CCCGATAAAC GATGCGATGT CCATCGCTTC CCCCGTGTCT TAGGTCTCGT   660

CGTTGAGGCA GCCGATGcTG CGCCGATAGG AGACAATTTT ATCGATAACT TCTTGCACAC   720

TTTCCCTCAC CACATAGCAC TTACCCGACA GCATTTGAAG CGTTACATCA GGTGTACAAC   780

GCATCGTTTC AATGTGGTGG GGATTTACCC AATTTTCATT TCCATTCAGT CGCGTCACTT   840

TAATCATCCC TCATCCCCAT CACGCCACCT GCCGCTTAAG ATACATTTTC ACACAGTCGA   900

CACATCAGCG CTTCAAACTC AACACCGTAT CCAACATGGT GTCTGATGTC TGAATCGTCT   960

TTGCGCCCGC CTGAAACCCT TTTTGGGTAA TGATCATATC CGTAAATTGA TCGGTTAAAT  1020

CTACGTTGCT CATCTCAAGT GTCCCTGCAA TCAACTTTCC CTTCCCCATC ACCCCCGACG  1080

TGCTAATGTT CGCTATCCCT GaGTTGTTCG ATTGTACGTA GGTGTTCTCT CCTGCCTTCT  1140

CAAGACCACC TTGATTTGCA AATCCTGCAA GTGCGAGCTG GCCAATGTCT TGGCTCACCC  1200

CATTTGAATA CACACCAGTG ATGACACCGC TTTGATCTAT TTTAAAATTT TCCAAATATC  1260

CCATCGCGTA ACCGTCCTGC CGGTAGCTTT GGTAGTACTG CGTTCAGCAA AcTGCGTAAT  1320

CGTATTGCGC GCGGTGCCAA TTTCACCCAA GTTGAGCGTG AAAGCGTGGC GCGTAACCTG  1380

cCCTGcATCG TCCGGaTTCG CACCGACAAC ATCGTACGAC GCTTCAAGGA GCACCTGTCC  1440

GGTAGGACCG GTCACGTTCC CTGCAGTGTC AGTCACTGAA GCGAGGTGTC CAAAATTATC  1500

AAAATTTACA ATAAAGGTGT TTGCCGCACC GTCAGATGTC CCCACCCCTA CACGCGTTTG  1560

CGTATCTACC TCTGTCCCCG GATCCACTGC GACAGTGGCC TGCCACTGAT TGTTCGTCCC  1620

CGGCACACGC GAAAAGTTAA TCTGCAACGT ATGCTGCTGC CCGAAGCTAT CATACACTTG  1680

AAAGTCAGTT GTCCACGTGG ACTTACGCAC GTCCGCTTCG TTCGCATCTG CAGCAAGCTC  1740
```

```
AGGCAGACGC TTGTCTAAAT TACAGGCATA GTGAACAGTG CTGGTCTGTG CGCATCTATC      1800

TTTTGCCCAA TGGGGATAAC GAGATCCTGC GTCTGTGCAG AGGAATTAAT TAAACGCTCC      1860

CCCGCCACGT CCTGCGCCAT CCAACCTTGA ACGCGCATAC CATTCGCAGG GTTCACGAGA      1920

GTGCCCGCAT TATCAACCCC AAAGGCACtG CGCGGGTGAA AAACGTCTTT TCCCCACTTT      1980

TCAGCACAAA AAAACCACTC CCCTGAATAG ACACATCCGT ATTGATACCC GTCGTTTGCA      2040

GTGCACCTTG CGTGTGAACA GTATCGATGC TTGCAATCAG CACGCCCAAT CCCACTTCCT      2100

TGGGATTCAC TCCTCCAACT TCTTCATTCG GACGCGCAGC cGcACTCAGT TGCTGAGAAA      2160

TAAGATCTTG AAAATTAACA CGCCCACGCT TAAAACCGGT AGTGTTAACG TTCGCGACGT      2220

TGTTCCCAAT GACATCCATG CGCGTTTGAT GATTCTGCAT ACCAGACACA CCTGAAAAAA      2280

GTGACCGCAT CATATGCTCT GTGTCCTCCT CATGTGACTC ATTTTCCTAA TCTCTCTTTC      2340

TCTGTTCACA AACCATACAA CTGCTACGAC GCACTCGGAT CTGCAATCAC CTTGACGTGC      2400

TCCCATTCGT ACCAGTGCGA CCCCACCCGC ACCTGGGGCT TGTCAGCACG GGTGACTGCA      2460

CTGATAAGCC CACGAACAGT GTTATCCGCC TCAGTGACTT CAACCATTTT TCCCACCGCC      2520

TGCAGCGCTT CAGTATTGCC AAACAGCGTT CCGAGCTTCT CTACCTGCGC ACTCATGTTG      2580

GCCATCTGCT CGAGCGAGGA AAATTGCGCC ATTTGCGCAA TAAACTGCGT GTCCTGCATA      2640

GGCGCAtAGG ATCCTGATGG GTAAGCTGCG CAATAAGGAG ATGCAAAAAA TCGTCCTTTC      2700

CTAACTCCCG CTTCGCACTG CGCGCGCCTG CCTCAAGCTG CTTGTTTATA ACGCGCACAT      2760

CCATTTCTAA ACGCGTACGC TCAGCGGCGG TCATTTCAAA CCGCATATTA GTGTTCTGTA      2820

CCATGCCCCG GCCCTCCTCT TTTTTCACCG GTTGTGTACG CTGACCCCCC TACAGGGTAG      2880

GCAAAACCGG CGCGCCTATT TAGGCAAACA CGTCAATCGT GAGCGCAnct CCcTGcGCAT      2940

GCCAATGGAC TTCTTGCACA ACAGGCTCAA CGCCCGCTCC CCCAAGACGC TGTGCAGCAG      3000

cgTAGGCTGC CGTCTGCGAT GCCAAATGCC CGTCCTCTGC GTGCGCACCA GCGCCAAACC      3060

ACTGCACATC AAACTGCGCA GskTCAAAAC CATTTGCCTC GAATGCACGC GCCAAATCCC      3120

CCAGATTTTC CTGAAAAGCT TCAAACGCCT CCTGAGAAGC AACGTGAATA GTACCCACCA      3180

CCCGCTTATT CTCCGACAGG GCAAGACGTA TGCTCACCGC ACCAAGGTGC TCTGGCTTCA      3240

GCGCAATGTC GATGTATCCG CGTCCGTGAT CGCGCAGCAC AACCCGTCCA GATTGCGCAA      3300

GCTCTGCACT ATGgCACGAA TGTGCGccGA AAGAGCCGCC TGAGTGGTAG CAAATCCTCG      3360

GATCCCCTGC GAnTGCGCTG TCTCCTCACG CGCGTGCGCA ACTCCTTCAA ACGCACGCTC      3420

CACGCCTGCA CGCTCCCCCT CACGCAGCGA CTCgTACCGT GCTCTTCCGC CCCCGCATCC      3480
```

```
ACGTCCGCAG CCGaCAGtGC GCCTGCGCCG CAGGAAGTAC GGCGCCCGCG TGCGAAGACA    3540

CACCCGACCC TGCACCGCGC GAATGCtACG CGCGTCCAGC ACAGTAAAGC GCGCGTCCGA    3600

AAAAGATCCC AACTCCTGAG GAGAATCCCC ATGCCGAATC TGCCCGTACG CAGTCGCCCC    3660

CTGGAAmGCC GCACCGGCGC TCACCCCTGC AGCAGCAGAG GCGTGAGACG CACCTGCCTT    3720

TCCTCCAAAG GAGCTGCCTG CTCCCCCAGA CTGCAAAACA GGCTCACCAC GGATAGCAGC    3780

CGCCACCAAC CGCTGCCGCA CCTCTGCATC GAAAATAACG TCAAAGACTT CAGACTCCGA    3840

CGATGCCGCG TACGTGGcTT CGCTCCCCTG cTCACGCAGC AGGGCAGGGG CAGCGCCAGA    3900

AGGAAGAGAC CCTCCGCGCA ACCCCCGCGC ACCAGCGCTT TCCCGCAGAT GCTCCCCAGA    3960

CTCCTGCCCA CACAGGTCCT GCACGTGCGG TGCAGTCTCC GGCACACGCT GCGsTaCGCG    4020

cgAGaCAGTC CTGCAGGACG CGCCGCACGC CCCCGTTCCC CGGATGAGGG CTGCTCTGGC    4080

ACAACAGACC GCGGCGTTAC GGAAGTTGCC TGCCGCTCTA CTACAAGAAA CTCCGGCGCG    4140

GACGGCCACG ACGCACCACC CGAAGCTTCC TGCGCCGCGC gnACGTAATC AAAGGGCCGC    4200

ACTCCTGACG CCGCCTCCTG CGCAGcaCGc AAACCAGTCT CGTACGCCAC GAGAAATACA    4260

TCCGAAAGAG ACTCTTCTGT GAAAACGTCC TGTTGGGTAC CCGAACCGGT TTGCTGCTCA    4320

TCAGCTTCCT CGCGCACCCT TTCGTGCACT GCAGCCGACG CAGGGAAAGA GAGACTCTCT    4380

GTGGCGCAcG CCACGCATGT GTTCATGCAA CGACTGAGCA AAAGAACACG GTGcTGCCGC    4440

ACACGACGTA CCGACTGAAA TAGTCTCCTG CGCAGCAGGC GcAGACTcCG CCACACCGAT    4500

ACCAATGGCC CGTGCCAGCA GTCCTCTCAG TTCCATGCAC TCCCCCCACT CCCTGCCATT    4560

CTCKGcGCAC TGCgCAGaCG TCTTAGAAAA AAGACCTGCC AGTCCGGTTC GGACGCATCT    4620

TTTCAATACA AAGCGsACGA GAAATTCAgC ACCyTCCcAA AGGsTcCAGC ACGCCGCTTC    4680

TTTCCCGTTC AGTATTTCCC CGTTTCCCCT GACAAAGAT GGATTCTCGG ATACTTTTCC    4740

CCTCCGTCTA TGGAAAAGGA AACAGCTGGC ATTCTCTGCT CCTATACGCT TTTTCACAGT    4800

GCGCTCGTGC TGGCGCTGTC CCTCGCGCAC GGGCGTACCC AGGTGCCCCC CAGCTCCACG    4860

CTCAGCTTTT TAACGGTCAT TGTACTCTGG CACTGTCTGC TCTTCTTTTT TCTTGTCGCG    4920

TATAGCAGAG AACCTGCAGA TACCACCGTG CCGTTTAAAC CGCTGCCTGA ACAGACAGCG    4980

CCTATTTGTG CCGCCGCATC TTCTGACTGT AAGGAGAACC GCACCGCGCT GAAAACGCTG    5040

AACACTGCAA CGCACATCAC GCTTATCCGT GCCAGTGCTA TTCCTATCGT TGGCTTTCTG    5100

CTTAAATTCC ACGCACTGGC GGGGCTTTCT TACTTCCTCG TTGCAGGACT GAGCGTTTTG    5160

TTCCTCACCG ATTTTATCGA TGGCAAAATT GCCCGCGCAA GACGAGAAAC GTCCCGCGTG    5220
```

```
GGAGAAACGC  TCGACGCAGC  AAGCGACTAC  GCGCTTATCG  GGCTCATCTC  AGCGCTTTAC       5280

TACCAAAGCG  GTGTGGTGCC  CCTGTGGTTC  TTTGTGCTTA  TCATCACCCG  GCTTTCGTTA       5340

CAAACGGTTA  TTGCCTGTGT  GTACGCGCTT  TTTGGCCACC  CGATGACCGG  TTCCACCGCG       5400

GGGGGCAAAG  CGACGGTGGC  CGTGACTATG  CTCCTGTACA  CGCTCGAACT  TGCCCGTCTC       5460

CTGCTGCCGA  ACCTTGCGCG  ATCAAACAGC  GGCGCGCGCT  TTTTTACCGG  GGCAGAAATC       5520

TTGCAGGATT  CGTCATTTTC  ACCGGGATAG  TGGAAAAACT  GTATCTTGGC  GTTCAGCATC       5580

GCCCAGGACG  CTCCCCGTAG  GAGAGACGAT  ACTTGCGCCG  TGCCTTGCAA  CACACAAAAC       5640

CTGTACCAAC  CGGGGCAAAA  GGAGTGCACG  CCCATGGATG  AAGGAAGAGA  AACTGTCCAG       5700

CCtgcGCATC  GCGCAAAGGA  GGAAAAAAAA  CAGGACGCCC  ATCTTGCATG  GGAGGTACGG       5760

AAACnGCACG  ArGCGTGCgC  CTGCGCGTTT  TTCACGTGCA  AGAACTCGAA  AGCGTTTCAC       5820

CGCGCAAAAC  GGTACtCGCT  TTGTAACGCT  CACTGCACCT  GAGTGGGTAA  TCGTCGTGCC       5880

GCACGTGATG  GAACGCGCAC  AACGCTTCTT  CGTTATGGTk  CGCCAGTGGC  gCTGCGGTTC       5940

ACAGACGGTG  TGTACTGAAT  TTCCCGGCGG  GGTTATCGAC  GCAGGGaGCA  CCCTGAGGCT       6000

GCAGCGCGCA  GGaGCTGTTT  GAAGAAACAG  GCAGACGCGC  TTCCTCTCTT  GCACACCTTG       6060

GCACCATACA  CCCGAATCCC  GCCGTGTTGG  AGAACCGCGT  GCACATCTTC  AGCGCCGAGT       6120

GTACGCCTGA  GnTACGTGAA  CCGCAGTTGG  ATACCGACGA  GTTTTTAGAG  CGGTGCGTGC       6180

TCCCCGTGCA  CGACGTGTAC  GAACGCATGG  GCCGCGCACC  CTTTGACCAC  GCGCTCATGG       6240

cGCAGCCCTC  TTTCTTTTTT  TGCGGGCGCA  TCCGCTTTCC  TCCCTGTAAC  TCAGTGCGGT       6300

ACGTCcCTGC  AGCGCGTCCA  TCTAGGTCGG  CATAGAGCGC  CGCTCTAAAG  GGGGGTATCA       6360

TCCCGGCTGC  ATAcTCTGCA  GCGCAGAGCG  TGTTGTGCAG  CAGCATCGCG  ATAGTCATCG       6420

GTCCTACTCC  CCCCGGAACA  GGCGTGATCG  CCTGCACCTT  GTGCGCCaTG  CGTCAAAATC       6480

CACATCACCA  CACAGTCTTC  TCCCGCGCGG  TGCAGTTGCA  TCTGGCACGT  GATGAATACC       6540

CACATCGATA  ACCACGGCGC  CGGTGCGCAC  AAACGGCGCG  CCAATGAAGC  GCGCCTTTCC       6600

CAGTGCTGCA  ACGAGGATAT  CTGCCTGCAC  ACAGATATCC  GCCAAACCGC  GCGTGTGACT       6660

GTGACAGAGC  GTCACGGTTG  CATCACAGCC  GGGAGAGGCA  AGGAGCACTG  CAAGCGGACG       6720

GCCAACGATG  GCAGAACGGC  CGACAATTAC  CACGCGTGCC  CCCGCAAGCG  GCACCTGCGC       6780

ACGCCGGAGC  AAGTGCACAA  TCCCCGCAGG  tGnCAGGGAA  CAAACCCAGG  CTGCGCAAGG       6840

AAGAGCGCAC  CACAGTTAAG  CGGATGAAAG  CCGTCGACAT  CTTTTTCTGG  CGCCACTGCG       6900

CGGCACACCC  TCGCTGCGTC  AAGATGCGCA  GtAACGGCAA  TTGGATCAAA  ATGCCGTGCA       6960
```

CCCGCGCGTC CTCATTGAGA CGAGCAATAA GTTCTAACAC CTGTGCGTGA GAGGCATGAG 7020

CAGGCAGCCG GTGCGTTTCC CCCCGCAGTG GGCGCGAgCA GGGGCACGCT GCTTTGCTGC 7080

AACGTAGTAC AAGAAGCCGG GTCATCCCCC ACCAGCACTG CGGGCAAGAA AnGGCGCCGT 7140

GCCTAcCGCC GCACGCAGCG CCTGCACACG CGTTGCAAGA CGgCCGTaCA CTCGTGTGCG 7200

GCTTGTTTTC CATCGATGAg GCGTGCGTCC ACGcGCCCAg TATAGACACG CGCACGCAAC 7260

AGCGCAAAGA CCGAACACGC ACGGACACTA GACGGAAGCC CAAGAAACAC CGTATGCTCG 7320

GCGTCGTATG AGCAGAACGT TCCGCGCGTG GCAGTGCGTT GGTGCGCTGT GTGCGCTCTC 7380

TCCCCTGCTG CCTGCCTACA rcTCCGAGGG CGTGCGAGAg GTACCCCCCT CCCAGTCTCC 7440

GCAGTGGTGG TGGCGTACGA GCCCATTCGC CCCGGGGATC AGCTGCTCAA AATTGGCATT 7500

GTTGCAGGCT GCCAGTTGTA CATAGCAGGG GGAAATGGAA CCAACGGCTC TTCGAGTTCC 7560

GGCACCAACG GTAACGGCAA CGGCAAACTG CTCGGGGGCG GGGGGTTTCA CCTCGGGTAC 7620

GAGTATTTTT TTACCAAAAA CTTTTCCCTC GGCGGGCAAG TTTCCTTTGA GTGTTACCGC 7680

ACGACCGGGT CAAACTATTA CTTTTCTGTT CCCATCACGG TAAACCCCAC GTACACGTTT 7740

GCCGTAGGcG ctGGCGCATA CCGCTCTCCC TGGGCGTTGG GCTCAACATT CAGTCCTATC 7800

TCAGCAAGAA GGCGCCGGGG CTTATTGCGG AAGCCAGCGC GGGGCTCTAC TACCAGTACA 7860

CCCCGGACTG GTCCATCGGC GGCATTGTTG CCTACACGCA GCTTGGGGAC ATTGCAAGCT 7920

CCCCCGACAA GTGCAGAGCC GTGGGCCTTG CCACCATTGA CTTTGGGGTG CGCTATCACT 7980

TTTAGCCCCG CCGCCGGGGC AGGTGGCGCG CGCGTCCCTA CTGGATAATG GCTTCAAGCG 8040

CAATTTCTAT CATTTGGGTA AAGGAGCGCT CCCGCTCCTG CGCGCTAGTT ACCGCGCCGG 8100

TTACCAGGTG GTCAGAGATA GTCAGAATGC TCAGCGCCTC GCGTCTGAAc TTTGCAGCAA 8160

GCGTGTACAG cTCCGCCGTT TCCATTTCCA CCGCTAACAC CCCATACCGG GCCCACAGGC 8220

GCCAGCTTCC TGATTCATCG TAAAAGACGT CAGAGGAAAT TACATTCCCC ACCTGCACCC 8280

CCGTGCCCAT TTCATCAGCA ACCGACACTG CCGTGCGCAG GAGCGACCAG CTTGCCGTGG 8340

GCGCAAAGTG CATGCCGCTA AACtGCGCGC GTTTATTGCA GAATCCGTTG CCGCACCCAG 8400

CGCACACACC ACCGATTTGA GCGCCACTTC CTCCTGCAAT CCACCGGCAG TCCCCACGCG 8460

GATTGCCTTT TGCACCCCAT AATCTTGAAA CAGCTCCGTT ACGTAAATTG AGTGCGACGG 8520

CAGCCCCATA CCTGTCCCCT GCACCGACAC GCGCACCCCC TTGTAGGTTC CCGTAAACCC 8580

GAGCATGCCA CGCACCTCAT TGTAGCAATA CGCATTGTGA AAAAAACGTC CGCCACAAAA 8640

CGCGCACGCA GCGGGTCACC GGGCAACAGC ACGCGCGGCG CAATATCCTC TCCCTTTGCT 8700

```
CCAAGgTGAA TACTCATCGT CACTCCCTCC CTTCGTGGGG CCTAGACCCA CAcGTTTCGG     8760

TAACCTCGCG CGTGTACGCT CAGCGCATGC AGCACCCGTT ACGCTTTTTG CGCAAGGTAC     8820

GCGTTTATAG ACGCCGCTGC ACGCCGCCCC TGCCCCATCG CACGAATAAC CGTTGCCGCT     8880

CCTAAGACAA TGTCTCCCCC AGCCCACACT CCCGGAATGC TCGTCCGTTG ATCCTCGTCC     8940

ACCACGATAG TACCCCGCTC GCTCACTGCA AGACTGCGCG TTGTCTTTGC CATGAGCGGA     9000

TTTGAACCAT TCCCAACGGC AACGATCACC GCGTCTGCAG CAAGTTTACA CTCAGCATCG     9060

CCGCAGGGCA GAAACACACG TTCTCCTGCA TCAATCTGTT CCTGACAATC GCGGAACACT     9120

ACCGCGCGCA CGTTCCCCTC TTCATCCCCC AAAATGCGGG TGGTCTGACA CAAAAAGTGA     9180

AACGTCACCC CCTCATCTTC TGCCTGTGCA ATTTCCTCCA CACAGGCGGT CATATCCGCA     9240

CGCGTTTTTC TGTACAGACA GTGCACCTGC TCAGCCCCTA AACGGAGCGC CGTACGCGAG     9300

GAATCTACCG CCACATTCCC TCCACCGACT ACCACCACTG ACTTTGCCGC ATACACCGGC     9360

GTGTCCGCAT GCGCAGTGTC ATACGCCTTC ATCAGCGTCG CACGCGTTAG GTAGTCGTTT     9420

GCTGCAAACA CCCCGCACAA TTCCTCACCC TCAATATTCA TAAAGCGCGG CAATCCCGCA     9480

CCGGTCCCGA TAAAAACTGC ATCAAAACCG TACTGCGAGA ACAGcTGTTC CAGCGTTGCT     9540

GTTCTGCCCA CCAAAAAGTT CATCCGGAAC GTcACCCCCA TTTTCTTGAG TGTTTCAATT     9600

TCCGTCACTA CCACTTCTTT CGGCAGGCGA AACTCAGGAA TACCATAGGT CACCACTCCA     9660

CCCGGTTTGT GGAGCGCTTC GAACACCGTT ACCGAATGGC CTGcACGCGC CGTATCTGAG     9720

GCAACTGcAA GACCTGCAGG CCCTGACCCG ATGACGGCCA CTTTCTTGTG CGTAGACGGC     9780

GCACAGTACG GAACTGTAAT TTGACCATGC TGCCGCTCCC AGTCAGCGAC AAAACGCTCA     9840

AGCGCACCAA TCGACACCGC CTTGGACACA TCCTTAAACA TCTTTCCCAC GGTACACTGC     9900

AATTGACACT GACGCTCATG CGGGCACACA CGACCGCAAA TTGCAGGGAG TAAACTCGTC     9960

GTCTTAATGA TATCAACTGC TTCCTTAAAG GCTCCCCTTT GGACACACGC AATAAACTCA    10020

GGAATCGGCA CTCCTACCGG ACAACCCTTT ACGCACGGCT TGGTTTTACA ATTCAAACAA    10080

CGCTGAGACT CAACCAGTGC CTGCTGCTCT GTAAAACCCA GCGCCGCCTC CTGCATGAGG    10140

AGCGACCGCT TTTTTGGCGG CAGCATACGC ATACGCTGCA AAGGGATCTG CGTGCGATCC    10200

TTCATCTTCA GCTCTTTACC CTGGAGCTGC GCCAGGCGCT GGCACGCTTC TTCCTGGAGC    10260

AGTGCGTGCG GCCGATACGT ACGCGCTTCT GGTTCTGACT CAACCGGTAC GTCACACGTC    10320

TTGGCATCGC TTACGACATT TTGTACAGAT GTCATACCTA CCTCCCCGCG TGGTGCTTCA    10380

TCTTACAGCA GTGGACATCA TGCGCTTCCC TTGCCTGAAA TGCCCTCATT CTCCGCATCA    10440
```

TGCTCTCAAA ATCAACTTGA T   10461

(2) INFORMATION FOR SEQ ID NO: 55:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13367 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

```
CTTCGCGCGC ATCGACATCC TCAATACCTT TATGGACAAG GCAGATACAG ATTCTGACGC      60

TTTCAGAGAA ATGTTCGACT ACTTTAACAC ATTTTTGCGT GCGTTTAGTG TCGTGGACGG     120

CAATGTAATT GCGGCTTACT TGGTGGTAAC GCGTGTTTCC ACGGTGCTGC CTCACCTAAA     180

TGCGTGTAGA CCCCATGGTT TTGCGGATTT GTACGCGCAT ATTGCGGATC CTCGATTGGT     240

GTACACAGAG ATAAAGGATA AGGGCCTCAA GTGGGAATTC GTGAATAGTG TGAAAAACTT     300

TGTGAGCAAT TGGAGCGATG AGTATGTCAA GCTGTTCCCC GAGGTGCTCT CTCTAGAGAT     360

TCTTCGCGCG CTTATGGAAG AGGGATATAA GGAAAAGGCA CTGAGGGTGG TCGAGGCTTG     420

CTTTGAATAC TATGCGGATA ATCGTGCGGC GGTTATTGGT TATTCAAGAC GGTAAnGGAT     480

GAGCCTTGGT TCCAGGGAGC TGCGCATTAC CGCAGAACAG CGGATTATCG TCCTCATCCA     540

CATTGTGGAC ATTACTTATC GGGAAATCGC TAACCGGCGG AACACCACTG AGAACCGAAA     600

ACTTAACnAG CAGGCTCTTT CGGTACTCTT TGGGAtGATC ATTTGCyAgA ACACyTCCAt     660

GCyTTCGCaC GATGTGGGAA CTACTACCCG TCTTTACACG TTATAAGTGA TATCCGGGGC     720

TTGATCCAAA GTTAAAGGTC CTTTGCGCCA TAAATTATTG AGAAGTACAG GATTTTAAGT     780

TTTTTGATAC TGAGGAACGT GTGGTTTCCG GACGTGGACT AGTGGTAACT GCAAAGATGC     840

TCAATGCAAA AAAGAAAGAA TTGCaGGATT TGCTTGATGT TCGTATTCCG GAAAATTCTC     900

GAGAGATTGG TAGGGCCTTA GAACTCGGTG ATTTGCGTGA GAACGCAGAG TATAAgGnTG     960

CGCGAGAAGA ACAAACAAGG TTGAACAATA TGGTGACTCG GCTACAAGAG GAGATTGAGC    1020

GGGCACAGGT ATTCGATCCT ACCACTGTTG TAGCTGGCAG AGTTTCGTTT GGTACGGTAA    1080

TTAGCTTAAA AAATCACACA AGTGGAGAAG ATGAGACATA CACTATTCTT GGTCCGTGGG    1140

AGTCGGCTCC AGAACGTGGT ATTATTTCGT ACATGTCTCC GTTAGGTAGC AATCTGCTCA    1200

ATCGTAAGAC AGGGGAACAA CTTGCCTTTA CGGTGGGAGA ACATGAAAAG GTGTATGAGA    1260

TCTTAAGCAT CTCTGCTGCA GAGATCTAGT GAGGAAGTGT GCGATGCGAA TTATGCGGaG    1320
```

```
ATTAATGTTA TTTCTTATGT GTCTATGTgc TGCGCTGTTT GCGCAAGAGC TGGTTCGCGA      1380

ACAGAGTGTT ACAAAGTCTG CAGATATTAC GGTGCTACTT GATACGTCTG GCACTATTTT      1440

ACCGTACCGT TCCGTGGTAA GCGGTAGTGT GCTAAAAGAT ATCGCTACTC GTTTTGTGCG      1500

TTTGGGTGAT TCGTTCCATA TTATTTCGTT TAGTGCCACG CCACGTCACG AGATTTCTCA      1560

GGTTATCCGT AGTGAGTTTG ATCTTTCTCA GGTAGTGTCT CGTTTCATGA TATTGCATCA      1620

GTTGGGGTTA TATTCTGACT TTTTAACAGC GCTAGATTTC GCGCGTACAC ACTTaCGCGC      1680

TTTGCCTGCA GCACATGAAA AAATTTTGAT TGTTGTGTCT GmCGGTATTT TTAACCCGCC      1740

TGCGCGTAGT TAgTgAAAAA CTACAACAAG GATCAGGTAA AAATTAACCT TGCACGGGCT      1800

GCCGCGGATC TGAGACGAGA GCAGGTGCGT GTGTTTTACA TAAAACTTCC CTTTCCCCAG      1860

GACATCCAGA TCCGCGATTT GGATGACAAT CTGCTGACTG ACCTACAAAA GACAGATGAT      1920

GTTCAAATCT CTGCAGTCGG TAGCTTTGCA GAAGGACAAA CAAGAAGGCC TAAGTTGGAC      1980

ACTGTGGGTG TGGTTTCCGA TCAAACGGGC GGCGTTGCAG ATAACCATGC AGTTGCTACG      2040

CACGGAAGGG AGGACGGGAC AGTCCAAGGG GTTGTTGGCA GCCATGTGGA GGTGGCACGC      2100

ACACAGGACA GACGCATAAT GCAGATCCTG CTAAAAGGGA AGGGGTTCGG CCTTCCTCAG      2160

AAGCAACTGA TGTTTCCCGC GAGTTCACGG AGGATTTGGG AATCAGGGTG AGTCCGGTTG      2220

ATTCAGATGG TTCTGTGCGT TTTTCCGAGA AGGAGCGCAC GCTTcCCGTG TTACACTTTC      2280

CAAGGGTCCT TGAGGTACAG GGTAAGTATG CAGAATGTAT GTTCGAGGTT GAAAATAGCA      2340

CGGATGCTCC CGTTTTGTTG CATTgGAGCG GGTGATTTTT GACAATGGCG TTGAGACTGA      2400

CATAGTTTCG GTGCAAACAG AGTCTTGTGC AGTAGCGTCC GGTGCACGCG CGATGTTGCG      2460

AACAACTTTT TTATTACCTA AGCGCTACCA CGAAGAGGGA ACGTACCAGG TGACCATGCG      2520

TGTACAGTTT GCAGATAACG TCCGCGTGTT CCCTCAGGTG GCAACAGCAG AGCTGCGCGT      2580

TTCTCCTTTG CCTTTTCTTG GATTGGTGCG GAGAGGTATA CATGGGGTTC TGTCTTCTGT      2640

AGGGCTTACG CATGCGTTTG GATATGTGTT GGACATGGTA GGGTTGAGTC GCACGGGTTT      2700

CGGTGCGGTG CTTTTGCCTC TGTTTGCTTT GGCTATCTTC TTAGTACTTG TATCAGCCGT      2760

GGTGTGTAGG TCAAAGCGCG TGTTGTCTCG TAAGTCATGG CGCGGAAGTC CCCGTACAGA      2820

GAATGGGTGT CAGGGTCCTG GTTCGATGTC TGATTTTCGG GCGCATTCTG TTAAGGAACA      2880

AAGGCAGGAT CAGGAGCGCG TGTATGCAGG CATGGAGAGA ATTGTATCTC AGCGTAAAAG      2940

CGATGTGCAG GATCGCCTCA GTGTATTGAA TGCGGCAACT GCATTTGGGC GTGATCGAGT      3000

TTCATTTTCC CCCAGGGTAA CGCGTGCGGA gCATGGATGT AGTCGGTCAG GAATGACTGA      3060
```

499

```
AATTTTTGTG TTTGATCAAA CACGTGCGAT TGGCAAGCGC AATATTCACG TAATGAAAGC    3120

AGGAACCCGT TTAGGGGTTG GCGGGCACAA GGGGGATGAC TTCCTAATTT TTTTGGTGCC    3180

GTTTCCAAGG CGGCTAGCAC AAGTGTATTT TGACGGTGAA GTATATCATC TTGCTATCTT    3240

GAAGCCGAGG TACTTCCCGT ACGAGGAGTC GAGTGTGGTG CCrAcTGCGT CGGCAGAGTG    3300

GTTACCCTTG TCTCTGACAG GGGGTATCAT GTGCCCTTCA CATTCCGCCA GTATGAGGAT    3360

CCCGCTGTGA GATTGAACAA TCTGCTCACC TCTATCGAAT ACGCTTGATC AAAGCGATAA    3420

AGGAACGAGT CGAGAGGTGG ATTCGGGACT CGATTGAGCA GATGGAAAGG GGGAAAGATG    3480

AAAACCGGGG CGGATCGCGC ACCGAAGCGT CCCCCTACCG TTGCTTCCTT CCGGACCTAG    3540

CGGGGTTTGG GGGATATTTC GTGGGCGGCC CCGGAACGGA GAGAGAGGGA TTCGAACCCT    3600

CGGTACCCTT TTGGGGCACA CACGACTTCC AATCGTGTAC TTCGGCCACT CGGACATCTC    3660

TCCTACGGCC GCACCCAGCC GGTTCTGAGA AGGGGGTGCG ACGTTTCCTC AGCCAACAAC    3720

GGAGAGAGAG GGATTCGAAC CCTCGGCGCC CTTGCAAGAG CGCTACGGTT TTCGAGACCG    3780

TCCGATTCGA CCGCTCTCGC ATCTCTCCTC AACAACAACG GCAGAGCCCC ACAGGACACC    3840

ACCCTCAGcG GGACAAGTCC CGTAATGAGA CTAGGCGGAT TCGAACCGTC GACCTTCAGA    3900

TCCGCAATCT GACACTCTAT CCAGCTGAGC TATAGTCTCA AGGGAGTGGG ATGCCAACCG    3960

GCCCCCAAAC CGGAGCaGGG GGGATTCGAA CCCCCGGCAC TCGGATGAAT GCAACTCtTA    4020

GCAGGGAGCC CGATTCGACC ACTCTCGCAC CGCTCCAAAA AACAGCAAAC AGACGCACcG    4080

TACCGAATAC TCCCCGCGGA GCAGGGGGGA TTCGAACCCC CGGTGCCTTG CGACACAGCG    4140

GTTTTCAAGA cCGTCGCCTT CAACCACTCG GCCACCACTC CGGACGCCCT TCCATCCTGC    4200

GTGTAAACGT TGCTCCTGTC AAGTCTTTGT ACGAGCAGCA TAAAAAAGTG GTACGTGTAG    4260

AAAACTTCCC TTCTGGGGAG AAGCTCTTAG AGAAGTAGCG TTTTTATGTT ACGCTCCCCC    4320

TTGTAGCTTG AGTAGGGGAG TATATGGACG ATGCAAGATA TGCAGAATGG AGTGCATCTT    4380

TGGTGCAGTT GCCCGATACG CATTTTTTTG ATCTTATGCG CCTCTATTTG GGTGTGCTTA    4440

AGACTCCATT TCATAAACAG AGGCTTGTTC AACAACTTAG TGCCTTCCTG CAAAGAAAGT    4500

CTATTCAGAA CGCTGTGGTG CAGATGCTTG ATGAACTCGA CTTGTTATTT ATTTCTGTTG    4560

TTATGTGCGT TCCCCGTGCA ACGCTCGAGC TGCTGACAAT TTTTTTTTAG AACGTGTTGC    4620

CCAGGCGGAG ATAAGAACAC GTCTACTGAA TTTAGAAGAA CGTCTTATTC TTTACCGCAT    4680

TCCTCAGATG CCTGGTGAGG TTACACAGGC AGAAGTCGCG AGCGTTGCGC AGAATGGTAG    4740

GGTGCGGCAA ACGCCGTGTT ATGGTATCAA TCCTCTGTTG CAAAAAGCAT gAGTACGGTA    4800
```

```
GCTGGACTCA ATCTTTTTCT CATTCCGCAA AAGCGGATGC GTCCATCCGC ACAGTTATTG    4860

ACAACAGATT TGATGCTGTG CGCATTGTAT tCGTTTTTTA CGCACGGGgA AAATTTATTA    4920

AAAGTCGAtG GGACGTTTAG GAAAAAGGCA TTtGTTATGT TCCAGGCATT GTTTCCtGTT    4980

GATCCGGATG TGGTGAGTGT GGCACTCCCT GCATATCTGC AGAGAGCAGG GGAGGAAAGG    5040

GGTACATCAC GTCTTTTACA GGAAGGTCGG CGCGTCTTGG AACATCTGGG ATTGATTGTC    5100

TGCGAATCAG CACAGGTGCA TGTGCAAGAT AAACGGTGGG CTTCTTTTTT CTCCTTAACT    5160

GCTCTGGAAC GTGCGGTGTA TTTGACAGTT GCCAGTACGG CTATTCTGCG CAAAGAGGTG    5220

CTCGTACAGC GAgCGCAGGC TTTGCGTACA CTTCTCTGTG TGTTGCACCC AGATGCGCAA    5280

TACGCACCTG AAGATCTAAC ACGCGTGTAT CGTATCTTGG TGGAAGAGGC AGCACCATCT    5340

GTTGCTGCTG ATTTTTTCTC TTCTTTGTCT TTGTCCAAAG ATACAATGCT GCAAAAGCGT    5400

AAAGGAGCTT TACATGATTC ATCGGTTTTT TCTATGCAGT CGGCGATCAC GGCTATACGC    5460

ACGGCCCAGC TTTTTGGGTT GTTGTGTGTG AAAGATGGAC TGTGCGCGTT GAATGAGGCT    5520

CTATTTAAAG GACAGTACAC GCGTGGGCCA GGAATGGTCT TGTCAGCGAC GGCAGAGTTA    5580

ACCATTTTCC CCGATGGAGA TATGCAaGGG GTTTTGCCAA TTTTATCCTG TGCGCATGTC    5640

TGCTCACTAC AAACAGTTGC CACGTTTGAG CTCAATAAAA AAAGCTGTAC CACTGGCTTT    5700

GCGCGCGGAT TAACAGTGCA GGCACTTGCA CAGGCTTTAG AATGTAAAAC AGGTGAGCAG    5760

GTGCCACAGA ATATACTATC TTCTTTCCGG CAGTGGTATG CaCAGATAAC CGCGTTGAcC    5820

TTAAGACGCg GCTTTGTCAT GCAGGTTGAT TCATCTCAGC AAGCTTTTTT TGAATCTGGC    5880

GGGCCACTGC ACCCGCTAGT GCGCACGCGT CTTGCAGAAG GAGTGTACTT TTTTGATGAA    5940

TGCCAAGAGT GTATGTTGTA TCaGGCcTCG CGCGAGCGCG TCTGTCCTAC CTGTGCGAGC    6000

CAATTGATAC AGCCACCCCG TTATTCCGCC CTGGTGAGCA GGGTGCACGT GCGCTCCATG    6060

TGCCTTCCTT TTCTTTTCCA GTGCGGTCTG CTCGGGGAGT CTCCGAGGAA TCAACGCGAG    6120

ATTTTGCACA TTTAGGTGCC TTTGTGTTGG AAACTCCGAA CGTTTCGTGC ACGCACAGTG    6180

CTGCAGATAC TCCGTCTATT TCAGAACAGA CCGGTGGGGT GGCTCACGTG CAGAGCGAAG    6240

AGGATGTAGA TCCGTCCACG TCTGGTGCAA CGGGTAAGTA TTGGGACAAG GCACAATGGC    6300

GCaAGGTGCa ACGGATGCGA CGTGCTGTGC GGCTGCAGCG GCTCAAAGAG TTTGAGGCGC    6360

ACCTGCAACA ACTAAAATTG GACGCAACAG AGCAGACGGA GCTACGTGCC CGCTTGCAAC    6420

GGGGGTTGAT TCTGGATAGA ATGCAACTTT CGTCCGAAAC GATCCGCaGG GAGAGAACGG    6480

AAGCGAGCGG GGTTGATTTT TTAGGCAAGT ATCGTCTTGC aGAGTGTGCG TTACGTTCTG    6540
```

GTGCTTTACT TGAGATTGAG ACTAGTTCAG GGCAGTCAGT GCATAAGATA GTGGGTACGG   6600

TGTGCGCAAT TGAAAAATGC GAAGAGGATG CGTTGCTTCA CGTGTGTGTA CACGCAGAAC   6660

TTCCCCCTGA GCGAGTATCG ATTGCGCGCG CGTCCAGGAT AGTGCTACTG AAAAATTCTA   6720

TTTTTTCTTG AGTCTGTTCT GAAGGGGATC CTTTTGTCTC TTGTAAAAAG GAATAGACGA   6780

GCGGGTAGGA TATGAGTCGT AGGAAACAGG GACGAGAGTT ATTCAACAGT CATGTGGGCG   6840

TGGTGTTGTC TTGTGTCGGT GCGGCAATGG GGCTTGCAAA CGTGTGGTTG TTCCCTGGAC   6900

GCCTGGTGGA ATTTGGTGGT GTGACGTTTT TAATTCCGTA TTTTATTTTT CTATTTGGTC   6960

TTTCCCGTTT TGGACTGATG GGGGAGTATG CTTTTGGAAA GACACTGCGC TGCGGTCCTG   7020

TGCGTGCGTT TACCCGTGTG TGTGAAACAC ATtCCATCGT GTTTTTTACG AGCACTACGA   7080

GGTAGCGGGT GGTTTCCGGT AGGAGTATTG CTCGCTACCT GCTCTTTTTA TGTAGTGATT   7140

ATAGGGTGGA TCTTGCGTTA TGTAGTATTT TCGTGCACGA ATGCACTTGC AGGTACTCAG   7200

GCGCACGACC TGTTTTACCA GGTTGCAGGG ACAAGTGCGA ATGTGCCGTG GACGCTTGCA   7260

GCTATCGCGC TCACAGCGTG TGTAGTGAGT GCGGGCGTGC AAAAGGGGGT GGAGCGAGGA   7320

AACATTATAA TGATGGTACT TTTTTACGGT GTCCTTGCGT TTATTACAGG ATATATATTT   7380

ACTCTTCCTA ACGCGTGGAT AGGTATGCGT AGAATGTTGG CATTTCAATC TTCATCATTG   7440

TGCAATCCGA GACTCTGGTT GTATGCATTA GGCATGTCGT TTTTTAGTCT CAGTTTGGGG   7500

GGCGCGGCTA TGGTTTTATA TGGCAGTTAC ATGCCAGATA CGGTGGACAT ACCGCGTACT   7560

GCATTTCAGA CAGCGACCTT AGATTTTTTG GCATCAGGTA TGTCCGCATT ATGTTTAATT   7620

CCGAGTGCGT GGGTTTTAGG TATGGACGTC AGCAGTGGAC CGGAGTTTTT GTTTGTAACA   7680

ATAACCCGTG TCGCCTCGCA GATACCGATG GGGGTGATGA TAAGTGTGnT AwTCtTTTTG   7740

TGTGTACTAT GTGCAGCGTT AAgTTCTGCA ATTGCTATGT TAGAAGTAAT ACTCGAGTCT   7800

TTTGTGCACA CGTGTACAGT GGGGCGCCGA ACGCTGACGT GGTCACTAGC ACTCGTGGTT   7860

GCGTTTGTAT CTCTTCCTCT GAATGCCTCG ATGAGAGTGT TCGAAACGTT TACAGATATA   7920

GTGGTGGTTA TACTATCTCC GTTATCTGCC CTTATGGGGA GCGTGATGAT ATTTTGGGTA   7980

TATGGTGCAG AGCGTTGCCG TGTAGCTATC AACCGGTGTG CACGCGGTCC GTTGGGTAAA   8040

TGGTTCACGC CGTATATGCG GTACGTGTAT TTGGGGCTTT GTGTAATGAT TATGGTGCTT   8100

GGGGTAATGT TCGGTGGTTT TTAGTGTGAT GACGCGCAAA AGCGGCCAAA CCCACAGTTG   8160

GGTAAATATA TTCTTTGCAA ATTGTCGACA CAACCGTTGA CGAGAGGGAT CGCAGGTGGA   8220

GAGGTGTCGT GCCGGGGATA TGTTGACACG TCCGTACTCC TCAGTTTGTG AGGCTCCAGT   8280

TATAGGAGGG GGGATAgCTA CGCGTGAAAA GATTTGCTCT TATTGGACTT GGAGACTTCG   8340

GTCTTAGCAT GCTAAAGGAG CTGCTCAAGC TCACTAACAA TATAGTCCTC CTGGACAGGG   8400

ATCGAACGCT CGTTGAAACC TACCGTAGCa GGGTGAGAAT CGTGCGCGCA ATTGAtGTGT   8460

TGGACGAATT CACTCTGTGC AAGATGATTC CACAGGaTAT CAACGCAGCG GTTATTGATC   8520

TGGGGGTTAA AATTGAATCA TCAATCATGA TAACAACGTT TTTAAAAAAA TTAGAAATTG   8580

CAGATATCGT AGTTAAGGCA TACAGCGCTG aACAAGGGCa TATCCtCTCG aGCGTTGGTG   8640

yTACGCACGT AGTkCTCCcG GACCGGGAGg CAGCTAAAAA AGTCACTCCT ATGATTGCTT   8700

TCGATCTTCT TTTCAACTTT ATGCCACTTT CTGCGCAgCT GGnCAATTGC GGAAATGGCT   8760

GTGCACGAGG ACTATGTGGG AAGAACTTTG CGTGAAGTGG ATGTGCGCAA AAACTTCTCT   8820

CTTAATATCA TTGCTATCCG TAAGCGCGAT GCAGAGGATT TTTGTTTTAT CAATGATCCT   8880

GAATACTGCT TTGAAGCGAA CGATGTGTTG CTCGTTGCCG GTTCTCACAA AGACATCTAT   8940

GCACAGTCGC AGGACAAGCT GGCACATACC CATAGCTTCA GCGACTTTTT CAAACAATGG   9000

TTCCTTACCA GCTGACTTCC CAATGTTCCG CGCACGGGAG TAGGCGCGTG TAATCTTCCC   9060

TTTTCCCGCA CATGCCTACG TAAAGGGGAA TATTTAGAGA GGGGGCTCAG CTTCAAGTTT   9120

TGAAAAATAA GGCTCAAGCG TTGCCGCTTC CCGAATTGAG GTTGCAGTGC TTACCACCGC   9180

AGCTTCAGCA CACGTGCGcT GCGCCCAGAG TAGTTGGGTA CACAAGTGCA TGTCTGTTAC   9240

CCGTGCGTCT AGAATATGCT CAATAGCCGC AATACGCTCT GTTTCTGTCG TTCCGTTGAG   9300

AAAGCGCAGG AAACTCGCAA AACTTTTATG CTCCGGTGTT TCAGGAGTTA CCGCCGTACT   9360

GTACGCTCCT ATGATTAGAC GTTCCATCGT CTTCTGGTTA AAGAGACTGG AGGTGCCCTT   9420

GTGATCGGCG TGCAGGTGCT CAATCGTCTT GAATATCACG TCAAGGGAGT GGAGCGGaTT   9480

TGGATCGCGG TAAGTTAGTG AGGAAAATAT GCCATGTACT GGATCTGGCA GGGTGAATGC   9540

ACCGTAAGCA CCACCTATCG TTCGAATTTT TTCCCAAAAC GGCTCAGTAC TTAGATATCG   9600

GGCAAACACC TGCTCTACCC CGCGTCTCTC CAAAGGAAGC CGTGGATGTG CAAGGGACAG   9660

CGCTGCAAAA CCCACTTGCA CAGGGGCTGG AAGCAGCGTC AnCATATTGC GGGTGCGCAT   9720

GTGCTGCAGG GCCTCTTGAA AGAGCACACC GTGAGCGGAG GGTATCTGTT GTTCGTGCGC   9780

TGTGGGCGCA GAGGTGTGGT GGATAAAATA AGTGGATAGC GGTGCACGAA AACACGCCAG   9840

AGGTTTTGCC AATGCATCTA GCGCTGTGTG TAACGATGTT TCTGTACCAC ATACACATCC   9900

GATCACTCCT GCAGTGAGCA GTTTTTCATG CAGCGCTTTG AGTTTGGCTG CCAGCGAAGG   9960

GGAGGCAACG GTTTCTGTAC ACTCTGTCCA CAAAGCACGT ACCAAGCGAA TCTGCGTGAC   10020

```
TCCAGTCCAG AGTTCTTCTA CGGCCTTTGC TGCATTCACG CGAGCGTTTG CCTTTGCAAG   10080

TGCAATGGAA TGTCCTGAAT GCATGGCAGC GCTATCCAAG TCATTTTTAT ATTGTGCAAG   10140

GATGTCTTTT AACCGTCGTG TATCTGTAAA AGAAAGACTG CGCACGTGTG CACACACGTA   10200

CGAAATCGCC TGTACGATAA AGCGCGACAG CATTTTTACG CTGACAACTA GCCACGCTCG   10260

CCCGACTATA TCGCTGCGTT GCAGTGTGTT CTGTCCCCTC AGTAAGGGGA GTATCTCGCT   10320

TCCCTGATCG CCTGCTACTA TACAGCGGGC GGCAAAGCCC CCGGTGAGAC GGGCAATCTC   10380

GGCAGATACC ACACTCCAAT GATGTGTCTC GGTTCCCATA CCTGTCAGTG CGTAGCCATA   10440

CAGTGGCAGA AGTTGTGCTT CTTTTACACT GAGCATATCT GCTGGGATTG CAAGGTGTAA   10500

GTACGTAATA TCGTTCGTGG CAAGCTCATG CACGAGAACA GGAACAGAAC CAAAAAACTG   10560

CATGGTTTCG CTCAGTTCTG GGGTGGGGAC TGGCAGTTGT TCCCGCTTTA TATGGGGGAG   10620

CAACGCAAGG AGTTCCTCCG GATCGGGTGT TGTCTGTCGT ACACGCAGTG ATTCTTGGTC   10680

AGCTCGGAGG CGCGCCGCTG CTGGCTGCGT GAGTGTACGG GAGAAATCCT GTACGTATTT   10740

TTCTAATTGC TCATCAAGTT TTTTTGAGAA GTCTGGGTCT GGGTGTACCG AAAGTACCGT   10800

GTACTGCGGG TTGCGCAgcA AGtGCGTGAG GATGAGATTT TCCACGTAGT GtGgATGGTG   10860

GTGTACCTTT TCACGCAGgc CTGCAGTGCG GGGATATAAC GCAAAGAACT TTCTGGACCT   10920

GCACCGTGCA ACCATCCACG CAGCGAACGC TGCATGAGCA CGAGAGAAAA AGGACCGTCA   10980

GAGCGGCGTA CTTCAGTATT TGAAAATTCG AGTGCATTCA GCGCTGTTTC CACTTCCTGT   11040

GGAGGGATGC CGTGCGCAAC AAGCGACTCT AGTGTTTCAA ACACGCATGC CTTTAGTGCA   11100

TCGACCTGTG TATGCTGCAC CCCAGTCATA CCTACAAAAA AAAGCATACG CTTTAGATCG   11160

ATGTGACTGC CGTTATATGC GTATAAATCC TCACCGAGTT CTGATTCTAA CAGTGCCTGT   11220

GCAAGGGGAG CAGCATCGTG ACCGAGCAAA ACGTGTTCGA GCAAAAACAC GTCCATTAAC   11280

TGTTCAGCCT TGTCTGATTC TGGGAGTAAC CAGCTGAGCA ATACGGCGCA CCGTGTTAAA   11340

TCCATCCCCT CGCTCGCCGG TGCGTACCCG GTGTACGTAC GGGGACTTTG GTATGCAGGG   11400

ATAGGGGGGA TGGGGGGCAA CGCTTTGCGG GCAGAAAATT TTGAAAGGCA TTTATCCTCA   11460

ATAAATGCCA TCTGTTTTTC GGTGGGTATA TTTCCGTACA GAAAAAGCTT GCAGTTTGAC   11520

GGGTGATAGT GTTTTTTGTG AAAAGCTTTA AACGATTCGT ACGTGAGACG AGGAATAACT   11580

GTTGGATGAC CTCCTGAATC GTGTGCATAC ACTGAGCCAC GTGTGGTCGC GTGTGTTGCG   11640

TGCTTATACA CAAGCGTATG AAAGTCTGCA TACACACCGC GCATTTCATT CAGTACAACG   11700

CCCTGGAGGG TAAGTTGGTT GTGCTCATTA AACTCAAAGC GGTGTCCTTC TTGCTTAAAG   11760
```

```
GTCCACTCTT CGATCAGGGG GAAAAAGACT GCGTCTGCAT ATACACTCAT AACATTGAAG    11820

TAGTCAGTCT CTACCAAGGA GGAGGCCGGA TATACTGTTT TGTCCGGAAA GGTTAGAGCG    11880

TTAAGAAACG TTTTCACGCT TTGTTTCGCG AGTATGAGGA ACGGATCCTT GAGGGGATAA    11940

TGCTGTGATC CACAGAGCAC CGAATGCTCA AGGATATGAG CAACCCCGGT ACTTGCTTCT    12000

TCTGCCGTCA TAAAACAGAA GGCAAACAAA TTCTCCGGGT CTTCGTTGAG AATGTGGTAC    12060

AACTCAAGCC CTGTTTTTTT GTGTCGAGCA TAGACACCCA CTGCCGAAAG CTCAGCGAGT    12120

GAATGGCGCC AGATAATTTC AAAACCGTGA AGAAGCGTAC TCATCGGTGA TTCTCACTCC    12180

TCTTCTTGCA AGCTATTTGG AAGCAATGTG CTGTTGCGCG CCGGCACTGC GCAATGTAGC    12240

TAAAAAGTGC TCAGTGATGG TGCGCGTATC ATGCGCTGAA GGCGGGAACG TGTCGTACTC    12300

ATCnCGCACA AGCTGTGCGC GGTTTTTAGA GAGATTAGAG AGAATTTTCT GAACAAAAGC    12360

AGGATGATTG GTGTTGATGA GCGCAGCAAG AGTTTTTTCA GAACAAGACG CCAGGTGTTT    12420

TTGCAAAAAT GTATCTGGGA GCGCGGGGAT ATCATCCAGC GTGAAAAGAT GCGTGCGGAC    12480

ACGCGCTGCA AGTGTTGGAT TTTTTTCTGC AAGGGCATGA AGAATTGAAT GCTCAGTCGC    12540

GCGCTCCATC TTTTTGAGAA TTGCCGCAAG CACTGCATGC CCGTCAAGAT CACGGCGCTG    12600

GGACAAATGG AGCGCTGCAA ACTTTTTGTG CAAGGAGTCA CTCATGACTT GCAGCACCTG    12660

AGGGTTAACG TGCTTTAACT TTGCAAGGCG AACGATCAAG TCCTTCTTCT CCTCTGTGCT    12720

GATATTACTC AAATAGTGCG CAGCGCTTTC TGGAGGCAGC TGCGAGAGGA TGAGTGTTTT    12780

GGTGGCAGGT AGTTCTCcTT CCAGGAGGGG GAGAAGTTGG GAGGCTTCAA GCGCAGCCAA    12840

AAACTCAAAA GGTTTCgGCT tGCCGCTGGC ACCGCCCGCT TCAAGATAAG ATCGGCCTTT    12900

TCTTCCCCAA ACGCTTTGGA AAGCATCGAc TGCGCAGCAC GCAGTCCACC GGTAACAGGC    12960

GACACACGAG CGCAGAGGGC AGAAAACTCC CGTAGGATCT CACGCGCTTC TTCTGGACTG    13020

AGGGGTTTGA GTGTCAGGAG CTCGGCAACC ACCGCCTCAA TCTGTGCAGG CTCAAGTTGC    13080

TTGAGCACCA GCGCCGCCTG CTCTTCTCCA ATGAGGGAGA GGAACTGGGC AATCTTTTTA    13140

TAAACGGTTC GGCCTCGGTC TTGTTCACGT ACGGTGGCTT TGATTAAGCC ACGnAGGAGA    13200

TTCGGTTCTA TTnCATAGCA AAGAGGACTC CGCGCGGTCT CCCGGCACAC GCAGTTGCAT    13260

TGTAGTGGAG GGTGTGCTCT TGACACAAGG GCGTGCAnAC CTTAAAAGGT GTCCCCCCCC    13320

CAGACGGGGT AGGGGTCCAA GGATGTGATG GCGTTGTCTT TCGGTTn             13367
```

(2) INFORMATION FOR SEQ ID NO: 56:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 6856 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

```
GCATTGcTGC GTCTCGATAG GCTGTTCGGT ATCAGCTGCG ATGATGAGGT GACCGGTCAG      60

TATCACTATG TGGTTATAGT TGGTGCGGCA GAGAAAAAGG TGGGGCTCAT GGTGGATGCG     120

CTGATTGGTG AGGAGGACGT ATCATCAAGC CACTGCGGGA TCAATTCACT AGTTCCCCTG     180

GTATTGCAGG GGCATCTATC CTGGGTGACG GTTCGGTGTC GTTGATTATC GATGTGGGGC     240

AGCTGCTTGA GCTTGGGTTG AAGCGGGAAA TATTGGCGCG TGAgcgTcGA GAAGCCACGG     300

TGTGGTAGGC GATCTGGGGC ACGGATTGGG GACTATGATA GAGCATATGG AAGCAGAGAT     360

CGGCATTCGG GAAAGTTTCG ACGGGGGCGT ACGTGAGCCG CTTGCGgTCA TAGACTTCAA     420

GATGGTTACC TTTTCCCTCG CGGGGAAGGA CTACGCGGTA GATATCATGC AGGTGAAGGA     480

AATTGCAAAG GCTGGGAGCT TTACCTATGT GCCCAATACG TCTCCGTTTG TTCTGGGGGT     540

GTATAACTTA CGGGGGGATA TTATTCCCAT AATTGATTTA AGGAGATTTT TTAATATTCC     600

CGCTCCGCGC AAGTCCCGGC AGGCGATCGA GAATATGGTG ATCGTCACAG TGGAAGATCA     660

GACATTCGGG GTTGTAGTAG ATGGCATCGA TAAGGTAATT GGGGTGTCAA AAACAACTAT     720

TCAGCCGCCA CACCCTATCT TTGGGGACAT CAACATAAAG TATATCCGGG GGGTGGTTGA     780

GGAGGCGGGA AAGCTGTACA TCCTACTTGA TGTGCACCGG ATTTTTTCCT TCCGTCTTGG     840

GGAGGAGGAA CGGACGGCAG TTGTCGATCG TGGTGTTGTG CCGTCTCCTT CACCTCCTGC     900

CGTATCTGTG CCGCCGGGGG ATGAAGAAAA TTTAAATGTT GGTTTCATTA GCGATACGTT     960

GGCCGCGTTT GGCCGTTTCT TTACCAGTGC AGTGAATGAG GGTTGGTTGC GCAgCCGGTA    1020

TCTTGTGTGG CGTGACGTGC GCTCTGGAGC TGAGGTACAG CTTCAGCATG AGGAGGATGT    1080

CGCCGAGTTC TTGAGTACAT TTCCTTCCCC GGACACAGGT GTGTTTTGGT CGGGGGAGTA    1140

TGCGGCGAGT GTGGGATCTG TTCTTTCTCG GATGCAGGTG GGAAAGGTGG TGACGGTGTG    1200

GAATATCGGT TGCGGTGCGG GTCACGAAAG TTACAGTCTT GCGGTGCTTC TCAGAAAAAC    1260

CTTCCCCGAC GCGGTGGTTC GGGTGCACGC AAGCGATTCG GATCTCTTCT CCATTTCCAA    1320

TGCTCCCATG cTCACTGTTC CTGAgCATGT GATCGGTGAT TGGTATAAGC CCTATGTGGT    1380

GAAGGGGGTG AGTGGTTCAT ACACCTTCTC CCAGGAAATT AAGGAGATGG TCCTGTTTGA    1440

GTACCACGAT TGTACGCATC CGAGTGCGCT TCCAGACGTC GATCTTATCG TGGCGCGGGA    1500
```

CGTACTGTCA TCTCTTGCGG TTCCAGTGCA GCACACCCTG TTGAAGGAGT TTTCTGAGAA   1560

GTTGAAGGCA ACAGGAGTTG TTCTGCTCGG TCAGAACGAG GTGATGCCTA AGGATACAGG   1620

ATGGTTGCGG CAGATTGAAG GCACCGTTGC GGTGTTCAGC AAGGAATAAT TAGCGCATGA   1680

GGAGTGGTGT ATGCGTGTAG AGTATATCAA CCCGTTCAGT GAGGCGGCGT ACGTGGTTCT   1740

GTCTGAGGTT TTAGCAGGGG AAACCAAGCG GGGGGACTTG TATTTGAAGT CTACGTGCAT   1800

GCCGGTGATG GGTGTTGCGG CTATCGTTGG CCTTGCAGGG GATGTAGAGG GGCGTGTGGT   1860

ATTTGACATG ACGCTCGATA CGGCGCTGAA GATTGCCTCT TCGATGAACG AGGAGAAGTT   1920

AGCGGCGTTT GATGAGcTTG CGCGTGCGAC GATCACCGAG CTCGCCAATC TGATCACCGC   1980

AAAGGCGGTT ACTACGTTGC ACGAGCTCGG ATTTAAGTTC GATCTTACCC CTCCGGCGCT   2040

GTTTACTGGG GACAACATGG AAATATCTAG TAGTGATATT GAAgCGCTTA TCGTGCCCAT   2100

GGAGACGCCT CAGGGTAAGG TGGAAATTAA TGTTGCCATC CGCGACAAAG TATAAGAGGG   2160

AGGAAGTATG ATTTCCAAGC AGGATTTTCC CACGATCAAC GATCGGGTTC CCGCAGaCaA   2220

AAACCGAATG GGGCGCCCTA TCGTGTGTTG GTGGTGGACG ACTCCATGTT CGTTTCAAAG   2280

CAGATTGGTC AAATCTTGAC AAGTGAAGGC TACGAGGTTG CAGATACTGC GGTGGACGGC   2340

GTTGATGGGG TTGAAAAGTA TAAGGCGATG AGTCCGGGCG TTGATTTGGT GACGATGGAT   2400

ATCACGATGC CCAAGATGGA CGGGATTACt GCGCTTGAGA AGATTCTTGA GTTTGATAAG   2460

AATGCAAAGG TAGTTATCAT TTCGGCGTTG GGGAAAGAGG AATTGGTGAA GAAGGCACTG   2520

TTACTGGGCG CGAAGAACTA TATTGTCAAG CCGCTCGATA GGAAAAAGGT GTTGGAGCGA   2580

ATTGCAAGCG TACTAAAGTG AGGGCGGATG TGTCCTGCGG GCTGTCTCGT ACGGTTTGCC   2640

CgCTTGCGTG TGTGGATGGT TTCTTGAGGT TTTTGCCTTC GCGCGCGGAG TGCCCGTCTC   2700

TCTGCGTGCG TGTTTTCTGT GTGTGTGCCG CAAGAGGAGA AgTGGTGTCT CCCTCAGCCC   2760

TTTGCTCGGG GCCCTGTGGT GCCTTTCCTG CGGTGTAGTT TCTATACTCC TTCGTAGTTC   2820

CTAGTTGGTT TGGTTGGAAA GGGTTCGGTT CGATTTTGAA GAGGTGCACA CGTTGTATTG   2880

TGCGCATGAA AGAGGGAGCG GTGTGTGCTC TTCCTGAAAA CGCTTGAGGT ATTTGGCTTT   2940

AAGTCGTTTG CAGATCGCGT TCGCGTTGAG TTTGCAGATG GCGTCACTGC GCTGTTGGGC   3000

CCAAACGGCT GTGGCAAAAG CAATGTCGTT GACGCCATAA AGTGGGTCCT CGGAGAGCAG   3060

TCCTCTAGGG CCTTGCGTGC CGACAGAATG GAAGACGTTA TATTCAACGG GACCGAGTCG   3120

CGTCGTTCGT TGAACGTTGC AGAAGCCTCT CTTACCGTTT GCGATGAAGC TGGTATCCTT   3180

TCGCTCGATG TGCCAGAGAT TTTAATTAAA CGCAGACTCT ATCGTTCCGG GGAAAGTGAG   3240

TACTTTCTTA ACGGGAATGC CGTCCGTCTA AAGGAGATCC GCGAGCTCTT TTGGGATACG   3300

GGAATAGGGA AGGTTGCGTA CTCCGTTATG GAGCaGGGGA AAATAGACCA GATTCTCTCA   3360

AATAAACCGG AGGAACGTCG CTACCTTTTT GAAGAAGCAG CaGGGGTGAC GCGCTTTAAA   3420

GTTCGTGGCG CGGAAGCAGC aCGGAAATTG GAGAAAACGG CGGAGAATTT GCGTCATCTT   3480

GAGGTTATTC TGCAAGAAGT AGAGAAGAGC TACGAGAGTT CAAAGCTCCA AGCTGCCCAG   3540

ACGCAACGTT ACCGCATGCT CAAAGAGGAG ATTTTTGCGC GAGATCGCGA TCTTGGTCTG   3600

TTGCGTCTGC GTGGGTTTTT AGAAAACCAA GCCCGAGCGG ATGGAGCACT CCAGCGCAaT   3660

cGCGCGGCGC GACGCGTTGC AAACACAGGT GGAGGAAGCA CAGCAGACGC TTTCTGCTCG   3720

CATAGGCGAG ATCAATGATA TGGAAAAGcg CGTTGACGCG CTCCAAAAGG AAATCTATGG   3780

CCTTGCAATT GAACAGAAAG CGAnCAAAAC GAGGCATCGC TACATCGTAA GCATCTTTCT   3840

GAACTGAAAG AGTCGATTGG TCAGATAGAA ATGCGCAAGA TTGGTGTAGA AAGTCGCGTG   3900

CAGAATTTGG AAGAAGAAGT AGCAGAGCAA GACGCACACG TGTATCAGTT AGGCAGTGCT   3960

CTATCCTCTG TTGAAGAGCA TATTGAATCG TTTGCGCGGA CTTGCACGTT GCAAGTGAGC   4020

ACGTCTCAGA GAATGATCAA ACGCTTCGCG ACATACAGGG ACAGATGCAA GAGATAAGTG   4080

CCGCGTGTGT TGAACTTGAA GCGTCCCTAC GTGACGTGGC AGAAGATATT GCCGCAGAGC   4140

TTGACACGCG CCTGAGTGCA GCCGGGTACT CTGCGCGCAA TCGGGCAGAG GCTGAGCGTA   4200

CGTTGGTAGC GGGGGTACAG CGCCTGCGAA CCTTCGTGGA GGGGAGAGCA CGTATTGTTT   4260

CAGACTTTCT GGTGGTAGAT ACCCACACTG AAGGGGAGCT GTGCCGGATG CTGACTACAG   4320

TTGTGGACGC GTTCAATGAG GCGGTAAAGA TAGTGCACTG CGTTGAGTCA GACATAGCAG   4380

AATATGCGCG TGTTTCTGCC CGGTTTATCG ATGAGTTTGT TGCTCCTCAG GGGATTATGA   4440

CCAAGAAACG TGAATTTGAG CGACAGCTTG AACAGCACCG TGCACAGCTT GAGCGGCaTG   4500

CTGCGCGTCA GCrCAaCTGn CAGGAAGAGA ACAAGCTCCT TGTTGGGAAG ATAGAAGCCT   4560

GTCGCAAAAC GCTTGAATCC CTGCGTGTGG ATCAGGCGCG TCTGCGTGCT GAAGCTGAGG   4620

CAGGACAAAA ACAGGCTGCA GGAACCAGAG GGGAGGTGGC ACGTCAGCGC GCAGTGATTA   4680

AAGAGCTCGA AGGGGAGTTG TTTACCGAGG GGGAGCGGGT GGCGgCGCTC GAAGAGCGCT   4740

TACTAGAGGT TGAAGGGGAA ATAGGACAGC TAGAACAGCG CGGTGTTTTG CTCACCAAAA   4800

GTCTTGAGAA CTGCGAAGGA GAGATCCGTG TGCGGAATGC CGCAGTAACA TCTGAAGAAC   4860

ATGCGCTCCA GGAAGCGCGC GTGGAACTTG CACAGGTGGG GCGGCAGCTT GAGCAGGCAC   4920

ATCGGGAGTT GATGCAGTGC GAAACTGAGA TTCGCAATTT ACGTGAACAT TTTCGAGAAC   4980

AGCACACCCG CGATCTGAGT GAGTTTGAGG ATTTAATACC GGGGaTTGAA AAAACGGCAA   5040

GTGATCTGCG CCaAGAGCGT GGGGAGCTTC aGGCTCGAGT GAAGGAAATC gGGGCgGTGA   5100

ACTTTATGGC GGTGGAGGAG TTTCAGGAGG TAAAGGAGCG CTACGAGTTT CTCGTTGCGC   5160

AGGTTGCGGA CCTTGAAAAG GCGCGCGCAG ATCTGCAGCG GGTAACCGAT AAAATTAAGG   5220

CTGAATCTGC AGAACTTTTC TTGGCAACAT ACCGACGGAT TCGTAAGAAT TTTcACGAGG   5280

TATTCCGTCG TCTGTTTGGG GGAGGTCGCG CAGAGATACG TCTTTCAGAT CCTGCAGCGG   5340

TGCTCTCGTG TGGAATTGAA ATCCTCGCGC AGCCACCGGG GAAGAAGCTC GAGCATATTG   5400

GCCTCCTTTC TGGTGGAGAA AAGGCAATGA CTGCAGTAGC GTTGCTCTTT GCAACGTATA   5460

TGGTGAAGCC TGcGCCGTTT TGTCTTTTGG ATGAAATCGA CGCAGCGTTG GATGAGCATA   5520

ATGTAGCTCG TTTTGTTGGG ATGCTTGATG AGTTTTCTGA CGTCAGTCAA TATATCGTAA   5580

TCACGCACAA TCGGCGGACG GTTTTGGGTG CACGCACCAT GCTTGGGGTA ACAATGGAAG   5640

AGCCGGGGGT ATCGAAAGTG GTTTCGATTG CACTTGAATC TGCTTCTGAG CGACCGGCTA   5700

ACGGCGAGGC AGGAGGAGCC ATTTGATGCG TCTGCGTGGG GTGGCAGGTG CCCTGTTGGG   5760

TGCGGTAGTG CTTGTGGCGT TGGGGCTGAT GGGCGTCTGG TGGGTGTTCT ATCCAAAAAA   5820

AGGGGACCGT GGGGCGGCTG TGGCTCGCGA GCCAGTGTTG TTGCACATAG ATCCTGCACA   5880

GATGGAGGCA GCTGATGAAC CGTTGACGCT TCCCCCTATC GAGCGTTCCC GTGAGCGGAT   5940

GTCGGCGTGG AGTGAGCAGG AGTGCCTCCG ACAGCTTGAG TATCCGACGG AAAAGGCGGT   6000

GCAGGCATTA GAGCACGCAA ACGAGAAACG TATACAGCAG ATGCTAGAGG CAGTACCGTG   6060

AGTGTGTGGG TGGCGCTCGC CTTGCTGGGA ATGTGTGTTT CGTGTACGCA CGTGCCTCCG   6120

CCTCGTGCCC TCATCGTTTC AAAGGAGCCG CCTCCAGCGT TGGATTCTGC GCCGCGCCCT   6180

GCGATTCCAG AAGCAGTTCC TCTTCCGTCC CCTGTGGAGG AAGAAATCGC CGGTCGCCTC   6240

CCTCCTGCAC CTGCCGCTGC ACCTGAGCGC GTTCCTGAGT CCTCACAGGA GCGGGAACAG   6300

AAACCTGAGT CTTCGAAGCC TCAGGTGGTA GAGCCGGTGT CGCTTGCCTC TCCGGTGAAG   6360

CCTCGCGAGG CTGGGAGTGT AcCTGATGTT CTTCCAGTAC CTGAAGTGTC GTCGCCGCAC   6420

GTTGCGCCGC CGGCACCCCC TGCGCCGAmA GCTCCCCGGC CGCATCGTCC CTCCCCTCCG   6480

CCTGTATCGC CTTCTGCATC CAAACCAAAG CAGCGCGCTG TACCTCCTTC TCCGCCCCCT   6540

GCATCAGAGC CTCCTCGTGA GGCGGAGGTG CAGGCTGAGC CTGAGCCGGC AGAGGATTCT   6600

CCACGCGCGA TGGTGCCTGA AGAAnCGACT GGAGGCATGA nGnnCCGCGC GTTTCGCnCG   6660

GATGACAGCT TGCATGGGGC AAAAACTTGA GGTTTTGTAT CCGGGGCGAA GTTGGGTGTA   6720

AGTGGGCGAG CATACTGCGC ACCTGGTTTG CGCTATCACC AGnGCAATTG GAGGAGTCGC   6780

ATTCGCTTTT TAACTTTATG CTGAGCGAGA GGGTGATTTT GTCTTAGnTT CTCCTAATTT   6840

GATGnGTTTC GGGGTG   6856

(2) INFORMATION FOR SEQ ID NO: 57:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 10928 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

CGCGTATGAA CGCAATGCCC AGGCGGTTAT TCCGTTGGAG CGTATCAGGC AGACAATCCG   60

TGCCGTTGAC GCGCGCGTGC AgtGCACTGG CTAGTTATTT TGAAAAGATA gGGGAAGAGA   120

AGCGGcTACG GGTCCTTGCT CGTCTACTCG AACGCTATGC ACCGCTTATC GGCGAGCAAA   180

AAATAACGGT ACgTTTCTTC GGTTATTGCG AGTCGCGGGT GCGTGATCTT CTCAATCAGG   240

CGCTTCCACG TGCTGTCCTG CGTTCTCTCA CCCCCTTTGA TAAGGCTGAG GCCTGGCGCG   300

CACAGTGCAG TGATGGGTTG ACTATTGAGA CGGAGGACGG GACGCTCCAG TGTCGGAGTA   360

CAATCGAGGA GATCTGCGCG CAACTTTTGT cTGAAAAGAG ACAGGAGTTG GCGTGTGCCC   420

TGTGCGGTAA TGGAGTGGTA GCGTGATCAA AGACGATGTG GTTACAGGCC GTGTAGTGAG   480

GGTGTCTGGT CCCATTGTGT ATGCCGAGGG CCTCTCTGCG TGCAGCgTAT ACGATGTTGT   540

CGACGTAGGg GAAGCATCGC TCATCGGAGA AATTATCCGG TTGGATGAGA GCAAGGCGgT   600

CGTGCAAGTA TACGAGGATG ACACAGGTAT GCGAGTCGGG GAGAAGGTGA CAAGCTTGCG   660

TCGACCACTC TCAGTCCGCT TAGGGCCTGG ATTAATCGGC ACCATTTATG ACGGTATTCA   720

GCGCCCACTT GAGCGCCTCT TCCAAGAAGA CGGCGCCTTC TTGCGTCCTG GTGCGCGTTC   780

ACAACCGCTT GATGGCTCCG TACGCTGGGA TTTTCGTCCT CATTGTAACG AGCGCGGTGA   840

GGCCCTGTGC GCGGGGATTC CGATTGCACC TGGGTCaGTG TTAGGGACCG TGCAGGAGAC   900

TCCTTCTGTT GTGCACACTA TCATGGTTCC TCCTGACATC CGGGGGAGCG TGCTATCTTC   960

GTTCAAGGGC GCAGGTGCTT ACACAATAGA TGAAGAAATT GGACGCACTG ATCTTGGTGA   1020

GCCGCTTTTT CTATCCCAGT ACTGGCCAGT GCGTCGTGCG CGTCCTTTCA GCAAAAAACT   1080

TGCAGTGTGT GAGCCACTAG TTACTGGACA GCGGGCGATT GATGTTTTCT TCCCCCTATC   1140

AAAGGGAGGA ACGGCGGCTA TTCCAGGGGG ATTTGGAACT GGGAAGACAA TGACGCAGCA   1200

```
TGCCGTTGCC AAGTGGTGTG ATGCAGATAT TATCGTGTAC ATCGGCTGCG GAGAGCGGGG    1260

CAACGAGATG ACAGACGTGC TCTCTGAATT TCCCAAACTC ATCGATCCGC GCACAGGACG    1320

CTCTCTTATG GAGCGGACGA TTTTGATCGC AAATACGTCC AATATGCCTG TGTCCGCACG    1380

CGAGGTGTCG CTGTATTCAG GGATTACCCT TGCGGAATAC TACCGTGATA TGGGTATGCa    1440

TGTGGCCATC ATGGCTGATT CTACCAGCCg CTGGGCGGAG GCGCTGCGTG AATTGTCTGG    1500

GCGCATGGAA GAAATGCcTG CGGAGGAGGG ATTCCCTGCG TACCTTCCGA CGCGTCTTGC    1560

AGAATTTTAT GAGCGCGCAG GACGCGTGGA AACCTGTGTG GCGCGCGAGG GCTCTGTGAG    1620

CATCATTGGT GCTGTTTCTC CCCTGGGTGG AGATTTCTCT GAGCCGGTGA CGCAGCACAC    1680

AAAGCGCTTC ATCCGTTGCT TTTGGGCCTT GGATCGTGAA CTTGCACACG CGCGTCATTA    1740

CCCTGCCATT GGGTGGATAG ATTCATACTC TGAATATGCG CAGGAAGTAA GTGCATGGTG    1800

GAGTAAGTAT GAcCCgCGCG CAGGCGTtGC GCGCCGCAGC CTTGGATTTG CTGAGAAAGG    1860

AACAGCgGTT ACAGCAAATT GTCaGGCTTG TCGGTCCtGA tGCGCTGCCt GGAGAAGATC    1920

GTCTGGTGCT AATGGTGTGT GAAATGATCA AAGGTGGCTT TCTGCAGCAG AACGCTTTTG    1980

ATCCGACGGA TGTGTTCTCC TGTCCCGAAA AGCAGGTGCA GATCTTGCGT ACCATAGTGG    2040

ATTTTCACGA ACGTGCCGTG GTGCTGCTGC GTGCAGGTAT TTCGCTTTCT GCGCTGTCCC    2100

AGCTTTCGTG CCGGGAGCTC ATCGTACGTA TGAAAAnTAC GTACGGGAAT GAGGATGTAC    2160

ACAAGATGCA GAAAGTGTAC GACACGATGT GCACTGAGTT TGACCAACTG AGTGTGTGTG    2220

CTGCCGCGCG CACACAAGGG GGGGAGAAAG TCGAATGAAG GGAGTGTGGT ATCGGGGTCT    2280

GTCCTCCATC GACGGTCCGA TCGTGGTGGC AAAGCGCCGG GAAGGTGCAT TCTATGGGGA    2340

GATTACGGCC ATCCGTGATC GCTTCGGTGC TCTGCGTACC GGCAGGATAA TTGATCTTTC    2400

TCAAGAGTGT TGTCTGATTC AGGTGTTTGG CTCCACGCTT GGGCTCAGCC TCGACGGTGC    2460

CTGCCtTGAG TTTTTGGACg TGCCGATGCA GCTGCGTGTC TGTGAGGGTT TGATGGGGCG    2520

GGTATTCGAT GGATTAGGGA GACCAATCGA TGGTTTCCCA GAGGTGCTCT CTTCTCAATT    2580

GCGTAATGTG AACGGCTATC CTATCAATCC GTACGCGCGC GTATATCCAC GTGACTTCAT    2640

TCAAACCGGT ATTTCTGCTA TCGATGGTAT GAATACGCTC ATTCGTGGGC AGAAACTGCC    2700

AATCTTCTCT GGGAACGGCC TTGCGCACAA CCGTTTAGCA GCGCAGATTA TCAGACAGGC    2760

AAAAATTCTT GGCACGGATG AGGCCTTTGT GATGGTATTC GCGGGTATGG GTATTAAGCA    2820

CGATGTGGCC CGCTTTTTTG TTTCTTCTTT TGAAGAAACA GGGGTACTGT CAAAGGTGGT    2880

GATGTTCCtG TCGCTTGCAG ATGCGCCATC TATCGAGCGT ATTATCACAC CACGCTGTGC    2940
```

```
ATTAACCGCA GCTGAGTATC TCGCCTTTGA AAAGAACAAG CATGTATTAG TCATTTTTAC    3000

AGACATGACA AACTACTGTG AGGCGCTGCG GGAAGTTTCC ACCACACGAG GGGAGGTACC    3060

CGGGCGTAAG GGTTATCCGG GTTACCTGTA TTCTGATTTG GCAGAACTGT ACGAGCGCGC    3120

AGGCAGAGTG AAAGGATCCT CCGGTTCGGT GACGCAGATT CCgAtCTTAA CTATGCCGAA    3180

CGACGATATT AGCCaTCCGA TCCCtGACCT GACCGGGTAC ATcACCGAAG GACAGATTGT    3240

GTTGCAACGC GACCTATCTC AGCGGGGCTT GTATCCGCCC ATTGGGTGTC TACCCAGCCT    3300

ATCTCGCTTA ATGAAAGATG GTATCGGGGA GGGTATGACA CGCGCAGATC ACCATGCGGT    3360

TTCAAGTCAG CTATTTGCTT CATACGCAAG AGTACAAAGC GTACGGAGCC TTGCCTCGAT    3420

TGTCGGAGAA GAGGAATTAC CTGCACTCGA TAAGTGTTAT CTGCGCTTTG GTGACTTGTT    3480

TGAGCAGTAC TTTCTCACGC AGgATGAGCA TGAAGATCGG AGTATCAGTC AGACGCTCGA    3540

TATCGGGTGG AGTTTGCTCT CACTTTTGCC GCGCACCGAG CTATATCGTA TCGACCCAAA    3600

GCTTATCGAT CAGTACCTGA CCGCTTCGTG CAGCGCGGTG AGTGATCAGT TGCGAAAGGC    3660

GATAGAGGAG GCCCGCACCC CGGTTGCGGA CGCGTAAAGA CCATGTGTCC TATAAGGCTC    3720

TTGGAGAAGG GTGATTTCTT TGCGGCGCTC CCTTGCTGTG TGTCTTGGCC ACGCAGGGAG    3780

AGGATACAGA GGTGAAAACA CCTTTAGCTC CCACCAAGTC GAATTTGGCG TATGTAAGAG    3840

ATCAGTTGGG TTTGGCTCGT GATGGTTATC GCTTGCTTGA GCAAAAACGA GAAATCCTCT    3900

TTATGGAGCT CACTTCTCTC TTGGAAGAGG TGCATCTTCT AGAGACTGAG CTTGATAAGC    3960

GTCGGAAGCA GGCGTATGCG TCGCTGTGGC AGCTGCTTCT TGCACAGGGC CGCGATGATA    4020

TTGCTGCCTG TGCGCTCGTA ACACCgGTGC CCTGCCGTGT GCAGCAGGAG GTGCTTTTAA    4080

TTGCTGGATT GCGATTTCTC CGTCTGGATG CAGTGATGCA GCCACCGAAG CTGCAGTATG    4140

CTGCGCTCGG CTCCAGCGCG TGCATGGATA GAGCGCGGGA GGACTTCGGG TTACTGTTGC    4200

AAACACTCAC GAGAATGGCA TCCGTACAGA CTATCGTATG GAGACTCGCG TCAGAAATGA    4260

GAAAAACACA GCGACGTGTG AATGCGCTGA GCAAGCAGAT AATCCCACAG ATGTGCGAGA    4320

CGTGCATGTA CATCGAAAGC GTGCTCGAGG AGCGCGATCG GGAAAGTACT TTTGTGCTCA    4380

AATCGCTAAA GGCGCGCAAG GATCCCACAA CCACCCTTTA GCACTCATCC GGCTGTACGT    4440

CCTGCGCTGC TGTTGTTCCG GGCCGACGCT ACCTCAGGGA GGCGCGTCCG ACACGCACTC    4500

TTCTTTTCCG CGGCCCTTTG CGTAGGtGCT CTTCTTCAGG AAAGcTGCGC GCGTGGGGGA    4560

CGTGCCGTGC TTCTCAGCGC GGCCCCTACG TTCTAGCAAA GCGGGaGCAA TGAGCTCAGC    4620

AATTTTTTTC GAAAGGGGAG AAACGGACAT TGCATACATA CGAGACGTGC GCACGATCTC    4680
```

```
CCCGGCTACA ATAAACTTCG GCCGGTCTTT ATACACACAG GACCCAGGAT GAATGTGAAT    4740

ACACTCTGCA GTGAGCGAGC GATACGAATC GCGGTGTTCT TGCACGCACA CGAACTGTAT    4800

CATGCCCGTG CCCACGCAGA TTAAGTATTC TTCCACGGTC CCGCCCGTTA GCAGAGGGAA    4860

TCCCATGCCA CTTACGATAA GCTCGAGCTG CTCCTTTATG TTAGAAATCT CAGCCATAAT    4920

CCGCTCGTCC AAATAGAAAT GTTGGCAAAA GGAATGTTTG TTATTTTCtT GCgcATACGC    4980

GCGGTAGATC CGCAAAAAAG ACACGAAATC CCCCATGGGA TCGGAGAACG TGCCGTGTGC    5040

CTTTTTTGTT TTTTCTTCCT GATCCTCTGA AAAGATTAGC GGGCTGCGAG CAGACAGAAA    5100

CGCCGCAGCG ATAAGTACAT CGTCAATAGA ATGGGGATAG CGCCGCAGCG CCTCTACAAT    5160

CATCCGGGAC TGCCgAGGAC CGAGCGGAAA CAGGCACATC ATTTTTCCAA TTTCACTCAG    5220

ACTCCGGTCA TCTTCTAACG CGCCGAGCAA GCGCAACGTG TCTTCTGCGC CGATAATACC    5280

ATGGGTGCCA GGAGGAGAAA TAAAATCAAA GTGTTCGAAA TCGTGGATAC CGAGTTCTGC    5340

CATGCGCATG ACTACCTCAG ATAGGTCAGT GCGGTAGATT TCTTCAAGGG TGTACGGTTC    5400

ACGCTGCTCA AAATCATCGC GCGAATATAG GCGATAGCAC GTGCCTGCGC GTACTCTGCC    5460

TGCGCGTCCA CGCCGCTGGT TACACGAAGC CTGAGAAATA GGAGTTTCGT CCAAACTTGC    5520

AGTATAGGAA AGCGGGTTAT ACGAATTTAA CTTCACCAAA CCAGAGTCAA TGACGGTAGT    5580

TACATCGTCA ATGGTGATGG ATGTTTCTGC AATATTCGTT GCGATGACGA CTTTTCTTTT    5640

TCCAAATGGC GCGCGGTTAA AAACTTGCTC TTGTTCTTCT TTACTCAATC TTCCATAGAG    5700

GGGCAAAAGA AAGAGCTTGC GGAACCAACG TTCATGGGAA AGACGGGTAA TACAATTTTT    5760

AATAGAACGC TCCCCTGGCA GAAAAATGAG TATGGCACCT TTGTCCCTTG AAGCGATAAC    5820

ACGCTCAACG ATACAAACGA TCTTTTCTAG CAAGGCGGCC TCCGCTTCCT TTGTATGAGT    5880

AGATGCaGGC GTATcAGGAG GATCGAAAAT AACAGTGACC GGGTATGCAA CCGCATCTAT    5940

TTTGATGAcA GGGCACTCAT TGAAATAGCG GGAAAACATG GCCGTGTTGA TTGTGGCAGA    6000

GGAGATGACG ATGCGGAAAT CATGCCGCTG TTGCAAGACG CGCTTAAGCA ATCCTAAAAT    6060

AAAATCAATG TTGAGACTCC GCTCATGCGC TTCATCTACC ATGATGATGG AGTATTTACT    6120

GAGGAGTGGG TCGAGCTTCA TTTCTTGCAG AAGGATTCCA TCAGTCATTA CTTTTATTTT    6180

TGTTTCGACA TCTGTGTGAT CCTCAAAGCG CATTTTGTAT CCGACAATGC CGGGCTGCAC    6240

GTGGAGCACC TGCTTGGCAA TGAACTCGCT TACAGAGAGG ACAGCAATTC TACGCGGCTG    6300

GGTGACGCCG ATAGCACCAC CTTCATTGTA TCCTGCTTCA TGAAGAATGA GTGGCAGCTG    6360

GGTAGTTTTC CCAGATCCGG TGGGGCTTTC GACAACAATG ACGTGATGGy kCGCGAcGCG    6420
```

CTAAGAATTT TGTCTTTCTG AGAGTAGACG GGCAACTGCT TGTAACTGAA CATGATTGCA   6480

AGCTCCTCTT ACTGCGTGTG GATAGGmCAG GATAGAAAAA AGAACCAGAA GTGGGAGTGG   6540

TGCGAACGGG CGGTAGGGAG CGTCCGCACc GCACTGCGgG AcgGTGcTGA GAGTACAGAA   6600

AGACGGAGCG ACCAAGCGCT AGTCATTGAC ACGTTCTTGA TATtcATtCG TCTCTGTATT   6660

TATCAGGATC TTCTCTCCTT GCTTGATAAA TAGGGGAACG CGCACGACAA GACCCGTTTC   6720

GGTAGTCACA GGCTTTGTGG CGCCAGAGAC GGTATCCCCC TTAAGATACG GCTCGCTGTG   6780

TGCAACACGG AAAACCATTT TGGTGGGAAT TTTTATGTCA ATGGACTCCC CGTTCCAAAT   6840

TAGGATGTCG TATTCGTCCC CTTCGCGCAA GTAGCGCTCT CTTCCTGGGA CATTCCCTTT   6900

GGAAACGAAA ATCTGTTCAA AACTGCGGGT ATCCATAAAG ACGAAGCATT CCCCGTCATC   6960

GTACTGATAC TGAGCGCGGT GGCTGTCTAC AACCGCATCT TCGACTGTAT CTGAGGTCTT   7020

AACTGTCTGA GTGAGCACAG AGCCGTCACG AAGATGTTTC ATTTTAACGC GCGCAAACGC   7080

AGCACCCTTA CCCGGGTTTA CGAACTCGCG CTCGACAACC AGGTACGGAG CACCTTTATG   7140

GAGCAGGACC GTCCCCTTTG CGATATCTCC CCCTCTAATC ATGTAATTCC TCTCTTATCT   7200

CCTAGTAAAC GTCTTGCACG ACCTGCGGGG GCGCAGTATA CCGCGCAGtA TATTTTTTAA   7260

AAGGCCTCGA ATGGAGGCAT TGACTTTTCG TCCCTTGCCT GGATACTAGG CGCCCTATGG   7320

CGAAGAACAC TGATATTGAG CACGACGCGC ATGAGCCGGC CGGGCACGGG GATGTGCGTG   7380

AGTCTGCCGT GGAGAATCCG TCTGCTTCGG CAGTGTCTGA CGGGGAGGAG CGCGCCACGT   7440

TTGCGCCGGA GtTGCTCCGC AAACCGATAC CGAATCAGCG CAAGGTGCAG CACAGGAGTC   7500

AGAGCCAGAG GTACAGCGCG CAGGAGAAGC TGAAAAGGGT GTACCAGAGA AGGCTAAGGC   7560

AGTAGTGCCG CTTGATGAGT TGTTGCCGCA GAAGGTCCAC TTAATTCCGC TCACCGGACG   7620

GCCTATCTAC CCGGGTATTT TTACTCCGCT TCTGATAAGC GATGAGGACG ATGTGCGTTC   7680

GGTGGAAAGT GCGTACAGCG ATAGTGGTTT TATTGGGTTG TGTTTGGTGA AAACCGACAC   7740

GCAAAACCCA ACTATCAGTG ATTTGTACGA GGTAGGATCG GTCGCTCGTA TTGTGAAGAA   7800

GATTAATCTG CCAGACGGTG GGTTAAATGT TTTTATTTCT ACACAAAAAC GTTTTCGCAT   7860

CCGCAAGCAC GTGCACCACA GCAAGCCTAT CGTAGCGGCA GTGCAGTACC TGTCCGATCT   7920

TATTGAGGGG GATCCACTCG AGATAAAGGC ACTTGTGCGT GGCCTTATTG GGGAAATGAA   7980

GGAGCTTTCT GAGAACAATC CACTTTTCTC AGAAGAAATG CGGCTGAATA TGATCAACAT   8040

TGATCACCCC GGCAAAATCG CCGATTTCAT CGCGAGTATC CTGAATATTT CAAAAGAAGA   8100

GCAGCAACGC ACGCTAGAGA TTCTGGATGt GCGCAAGCGC ATGGAGGAAG TCTTTGTATA   8160

```
TATCAAAAAA GAAAAAGACT TATTAGAAAT CCAGAGAAAA ATTCAAAATG ATTTGAACAG   8220

TCGGGTGGAG AAAAACCAAC GCGAGTATTT TCTGCGTGAA GAGCTGCGTT CCATCAAGGA   8280

AGAGCTGGGT CTTACCACCG ATCCAAAGGA GCGTGATCAG CGGAAGTTCC GTGCGCTAAT   8340

AGATTCGTTT CACTTTGAAG GGGAAGTGAA AGAGGCTGTG GAGAGCGAAT TGGAAAAGCT   8400

CTCCCTTACA GACCCGAATT CCCCTGAATA TTCaGTGGGT CGAACGTACC TCGAGACGGT   8460

GCTCTCTTTa CCTTGGcACG CTCCTGAGAA GGAGGAATaT GACTTAAAGA AAGCTCAGAA   8520

ACTGCTTGAT GAAGACCATT ATGGACTCGA GAATGTCAAA GAACGGATCG TGGAGTATTT   8580

GGCGGTGCGA AAGTTACGCG CCGATACCAA AGGCTCTATC ATCCTGCTGG TAGGTCCGCC   8640

GGGTGTGGGA AAAACCTCGG TGGGCAAGTC GATAGCGCGC GCCATCCACA AGCCCTTCTT   8700

CCGTTTCTCG GTTGGAGGGA TAAGCGATGA GGCCGAAATC AAGGGGCACA GACGTACTTA   8760

TATCGGCGCC CTGCCGGGTA AGGTGCTACA GGGGCTGAAA ATAGTAAAAA CTAAGGCTCC   8820

CGTGTTTATG ATCGACGAGG TGGACAAGAT TGGTTCTGGC GCGCGCGGCG ATCCTGCGGG   8880

GGCTCTGCTG GAGGTGCTTG ATCCGGAGCA GAACaCTACG TTCCGCGATC ATTACTTAGA   8940

TTTGCCCTTT GATCTCTCTC ATATCGTGTT CGTGCTCACT GCCAATAGCA CCGATCCTAT   9000

TCCCCGTCCA CTGCTGGATC GCGCTGAGAT TATCCGTCTT TCCGGTTATA TCGATACGGA   9060

AAAGGTTGAG ATCGCAAAGC GCCATCTGGT GCCAAAAACG CTGGAGAAGA ATGGTTTAAA   9120

GCGTGCGTGC GTCTCTTATC GGAAGGAGGT GTTGCTACAC CTGGTCCATT CTTATGCGCG   9180

GGAGTCTGGG GTACGGGGGC TAGAAAAAAG CCTTGACAAG CTGCATCGCA AGCTTGCCAC   9240

CGAGATCGTG TTAGGGAAGC GATCGTTTGA TGACAAGTGT TTGATGGATG AAGCTCTCAT   9300

AGGGACCTTT TTAGGGAAGC CCGTGTTCCG CGATGATATG CTCAAAGACG CGAACAAAGT   9360

TGGTACTGCG GTGGGTTTAG CCTGGACTGG CATGGGGGGA GACACGCTCC TTGTTGAGGC   9420

AATTACTATA CCAGGAAAAG CAAGTTTTAA GCTCACTGGG CAGATGGGAG CGGTTATGAA   9480

GGAATCCGCT TCTATTGCCT TGTCCcTGtG CGCCGTTACA GCGCGCAgCA GCGTATCnTT   9540

CGCCGAATTG GTTTGAAAAG CGCGCAATAC ATCTGCATAT CCCCGAGGGC GCAACCCCAA   9600

AGGACGGTCC GTCCGCGGGG ATTACCATGA CCACCACGCT CTTcTCGTTG CTCACCCAGC   9660

AGAAAGTAAA GCCTCGCCTA GCGATGACTG GAGAACTCTC ACTGACCGGA CAGGTGCTCC   9720

CCATCGGGGG ATTGAAGGAA AAGACTATCG CsCACGGCGC GGTGGTATCA AGGAGATCAT   9780

CATGCCAAAA GCGAATGTGC GGGATCTGGA CGAAATCCCC GAGCACGTCA AGAAGGGCAT   9840

gTGTTCCACC TAGTTGAATC GATGGAAGAG GTCCTTTCTC TCGCCTTCCC CAAGGGGAAG   9900
```

CGTGTCCGTG CTGGCACTGC CGCCCAATCT GCTTCTCCTG AAAcCCTTAC AGGCTGACGT   9960

ATGCGCTTTC GTGCACGCGT ATCTCAGTCA ACTGCGAAgT GcGTCGTGTT CACAGGAGGC   10020

GGCACGGgAG GACACATTTT CCCGGGAATT GCAGTTTTTC AAGCgCTTGC gCAcrGGCGG   10080

cGGtGCGTGT CGTGTGGATT GGTGCAGCGC GTGGTGCTGA TCGCTCCATA GTGGAATCTG   10140

CCGGATTAGA GTTTTGTGGT ATCACCGCTG GCAAGTGGCG TCGGTACGCG AGTGTGCGCA   10200

ATTTTTTTGA TGTATTTCGA GTGCTCGTCG GTACGGTGCA ATCCTATTGT ATCTTGCGCG   10260

CTTTGCnCCC GCAGGCACTA TTTTCTAAGG GAGGGTTTGT GTCCGTGCCG CCGTGCATCG   10320

CAGCGTGGCT TTTGCGCATA CCCGTTGTCA CGCATGAATC GGATATCAGT CCAGGACTTG   10380

CCACACGCAT CAATGCGCGT TTCGCCGATC GTATTTTAGT CTCTTATCCG CACACGTCCT   10440

GTTATTTTCC CCgTGcGCGA CGCgcAGCAG TTCACTGCAC GGGGAATCCT GTGCGACAAG   10500

ATTTTTTTTC TGCACAGGCA GAGCGTGCAT ACCAGTTTTT ACGCATTGAC CAAAAAAAGC   10560

CATTGCTCAC AGTCCTCGGA GGAAGTAGCG GTGCGCGTGA CCTAAACGCG CGTGTTCTTT   10620

CATGTAGCAC CTTCCTTACC GAACGCTTCT ATCTTGTCCA TCAATTTGGC GCAGgCAACG   10680

AGGACCAAAT GCATACTATC ACCAATTCGC TTAGcGTCAA TGCTCGGCAT GCCTACATGT   10740

CGTTTCCTTT CATTCAGGgC ACATCTGCCC GATATACTCG CCGCGAGCGC ACTGGTACTC   10800

TCTCGTGgCT GGTGCGAACG CGGTGTGGGA GTGCGCATGC TCGGTAAACC AATGGATTCT   10860

TTTTCCTCTC GAACGAGGGA GTTCCCGTGG GGATCAGATT GAAAATGGCA GAATATTTTA   10920

GCGCACAC   10928

(2) INFORMATION FOR SEQ ID NO: 58:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 3237 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

TACAACGCCG TGAGCAACCC TGCACTCAAA AAGATCAAAG CAATAACATG GGGAGGAGCG   60

CCCATCGCTT TTAACATCGC AATTTCTTTA CGCCGTTCTG TCATTAGCAC CACCAGTACC   120

GAAGAGATGT GCACTGACGC TACTAGCACT ATCAAATACA TAATGAATAA CAATAACTTT   180

CGTGATGTTC GGAAAGAATG AAATTGCGAT CGATTCATGT CTTGCCACGT GTACGCACTA   240

AAATGGCTCG GTAATTGTTC GTGTACTTGC TCAATAAAAC GAGTCATCGC CTTAGCGTCA   300

```
AACGCGTCGG CAGTTTTTAC CACAAAAGAA AGTAGCGCAG ATGCGGGGGA GAGAATTTTC    360

ATTCCCAGCG TGAGGGGGAT AAATACCCAC AACGCATCAA GCTCCTGATA TCCGCAAGAA    420

ACAATACCTC CTACCACCGC GCGCACCATT TTGGGGACCG CACGACCTGT CCCTCCTTGC    480

ACGAGGGTGA GTATCTGGCA CGTGTCCCCA CAGCGCACCC CAATGCGCTC AGCGATGCGT    540

TTTCCCAATA TTAACGTGTG TACTCCtGCG GCCTATCTAC CAGTTCAAGT GAACCTTCGA    600

CGGTTAAAAA TGGACGAAGT CcACGcTCAC TAGAAAAAAA ATCAGGGGGA ACTGCGCGGA    660

TATTCCCCCC TGCACGCCCT GTTTTTCCGA TTACAATACC ATCTCCCTGA AGGTGCATCC    720

ATCGTGAGTG ACAGTATGGG CCAAAGTCcT GCGCCATAAA TGCATTAAAT ATGCgctGCG    780

CGTCTTCATA TCGTTGCGTT GCCGTTTCAT TGGGAGCAAG CGGCAGTATA TCGATAAACT    840

GGAGGTGACC CGATCCGAGT TCAATCATCC GTGTGGTGAT CCCTTCAATC ATTCCATCAG    900

ACACCACAAG GACAACAATG AGTGGGATGA TGCTAATCCC GATGCCGAGC GCGGCACAGA    960

AAAAACTTTT GCGCAAAAAG GAACGTCGTT TTCCTGCTAC CGGAGTACCT GATAGAAAAG   1020

GTACTGGcGT AGGCAGTACG TGGTGCGCAT CACCGTGTAG AGATGGGGTG TGCCCATATC   1080

CTGcGCACAC TCCTGCGCAA CGTAATGCAC ACATGAAAAT AACTCGAATC AGATTCACCG   1140

CGGCGCACCT TTGTTTCATA TGCGTATCAA ACTTCCCTGC TGTAGCTGGT AACGGTAtCG   1200

GTCATCGATG CaATACGTGG GTCGTGCGTT ACAATGAGTA ACGTCTTTTG ATATTCCTCT   1260

GTCAGAGAGA ACAGCAGATC CTGCACTATC AAAGCGTTCT TGGGATCCAA ATTGCCAGTC   1320

GGTTCGTCCG CAAGAATTAG GGTGGGATCA TTGATCAGTG CACGCGCAAC TGCTGTCCGC   1380

TGTCTTTCTC CTCCTGACAT TTGTGCAGGA AAATGATGGG CGCGCTGCAC TACGCGTACT   1440

TTTTCTAGCA ATTCGTATGC GCGTGCACGC ACcTCACGGT AACTTTTTCC TGCGATAAGT   1500

CCAGGCAACA TGACATTTTC AAGCGCAGTA AAATCCCTCA GTAGATGATG AAATTGAAAA   1560

ACTAATCCTA AAAACTGTCT GCGGTATTCT GTCAGTGCGT GCTCATGCAA AGTGAGTACG   1620

TCGCATGAAA GCACTCTGAC GATCCCCGAA TCAGCGTGTT CCATTCCTCC AATAATATTC   1680

AGTAAGGTAC TTTTACCGCA GCCGGATTCT CCGGTGATTG CAACCTTCAC TGCACGCGGC   1740

ACGcTAAATG ATACGTCAGA CAAAATCTGT ATACGTTCTG TTGCGCAgCa GAAnGCTTTT   1800

ACTTACTTGT TCGACAGAAA GAATTGGGTC AtTCATCGCG TAGCACCTCA GCCgGCTTGA   1860

GCaGGAGTAT TTTACGCGTG GCAAGGTACG TTGCAACAGA CGCAGAACCT GTGCCAAACA   1920

GAAATACAAA CAGTACCTCC TGAAAGAAAA TCTGCACGGG AATACGCTCC ACGTTGTAAA   1980

AATATTGCGT ACCAAACACA CTGAAAGAAG GGGTTTTCGT TCCCGAGAAG AGGGAGAACA   2040
```

GGAAAAACGC AGAATTTACA GCAGTCTCAA TGCACGCAAT TATTTCGTTA ACGTGGATAG   2100

TAATGAGCAA TCCCAGGAGT ACCCCCAAGA GAGAGCCCAA AAAGCCAATC ATAATGCCAT   2160

TGCCGATGAA CAGAATCTGC ACGTGACTGA CAGGGGCGCC AAGTGAAACG AGCATAGCAA   2220

TTTCTTCCTT TCGAGTGCGA ATAGAGCGGC GCATGCTGTG ATAAATGTTT ACGGTTACCA   2280

CCwTAAAAtC AAAtGACAAG AAGTATCATG ACGTTCTTCT CTATGCGGAG CGCACTAAAA   2340

AAaGCaCGGT TGTACTCCCG cCAGGATTCT GCCttGAGAT scAGGAATGT GTTGTGCaAG   2400

AAAGAAAAGG TAGCGATCGT CTCGCTCATG GTTATTTAGT TTGACTGCCG CGGTAATATC   2460

aGGCGTCGTA CCAAATAAAG TGGTGCCCAT GTCCAGAGGA ATGTACGCAA ACGTGGAATC   2520

TACTTCGTGG TATCCCGATT TGAAAATGCC CGTTACCGTA AGTTATTCC AGCCTGGCAT   2580

TATCTTTTGT GTATCACTTC CTGACAGGGC AAGCGTGTCA ACCTGATCTC CGGTACGTAC   2640

CGAAAGGTGG CGCGCCAGTT CATATCCGAG CACAATGGAG TGCTTTTTAC TCAAATTAAA   2700

ACTTCCGGAT GTTATCGGGA GTGCACGCGC CAGCAACCTA TCCCGATGGA AGATATCTGC   2760

AGGAACTGCA CGCACAAGCG CACCGTGTTG CCGATAATAG TTGCCTTGCA ATAAGGCATG   2820

CGCTTCTATA AATGGATAAA AGGATTGATA GcCGCCTAAC GTCTCTGCAC GTTTTAtGCG   2880

TCAACACTGC CATATACACG AACGTGTGCA GAACTCACCT GTAAAATGGT GCCAATAAAA   2940

CCCTGCTGGA AGCCGTTCAT AACCGAAAGG ATGACAATTA AGGTAAGTGC CCCAAAGGCA   3000

ATGCCTAATA TAAAAAAAAG ACTGGTAATC GCGTTyGCAC TCCGCGCGCG CACTGAATTT   3060

AATCTGCGCA CCATAAAACA CATCCACCGc AGCGTTTGCA CGTGGGTGTT ACTCATCGTG   3120

TACTTCCTTt GTAGAGTAAA TTTCCTTCCC ACGCTCAAAT ACCTGCACGT CCTGCGCACC   3180

TTCCTTTTGG TGAATTATTG CCTGTAATAC nCCGTTGCGA TAACGTTTTT CTAACAC   3237

(2) INFORMATION FOR SEQ ID NO: 59:

       (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 2582 base pairs
           (B) TYPE: nucleic acid
           (C) STRANDEDNESS: double
           (D) TOPOLOGY: linear

       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GTCGTATCCG nGnTAGTCCA CCGGTTCCTG AAAACACCTG CTGCGCTGCA CGGACACCAC   60

CTTTCCCCCA AGTTCATCCA AAACAGGGTC GCAAACCGCT GCGAGTTGAA AAGCTCGACG   120

CGTTGCTGCA CACGTAATCT CATAAATATT GCCCTGAAAC TCAATCAACT GCTGCATTGG   180

GAAAATCATG GGTCATCGTA ACACCTATCG CGCCGCGCAA CAAGCACTGC TTGACGTACA   240

CGATCCCCCC TGTGTAGAAT GCGCACCGCG GGCAATGGCG CAGTAGGTTA GCGTACAAGT   300

CTGGGGGACT TGGGGTCCCC GGTTCGAGTC CGGGTTGCCC GAGGAACAGC TGGCCAGCTC   360

ACCCCGGTAG GCCGTTTGTG TCTTTGGGAA GCGGCGTGCC CGGCTGCGTC TCGCCTGTAG   420

GTTGTAGCTT CCGTAGTGCG TTCTTAGGCA GGTGTGTAAG GAATGGCGGG GCTAGTAGGC   480

ACTCAGGGTC TATCGGCGTG CTATTTATCC GTGCTTCCCA GTGTACGTGA GGCCCGGTGG   540

AGAACCCGGT GGTTCCGCTG CGTGCAATGA GCGTGCCCGT TGACACGTAC GTATCTTTTT   600

TCACCAGCAG CTCGTTCAGA TGGTAGTACG CTGTGTACAG CCCCGGGGCG TGCTCCAGTA   660

CCAcGCTCCA AmCCGTGGTT GTCCGTTGCT CTGCGAGTAA CmACCtCCCT GTACnTGcTG   720

cATACAmCGC CGTTCCCACc GGAACTCCAA AGTCCTTCCC CCAGTGGTAC CTGnCAGAGC   780

GCGTCCCGTC GGTGTACACA AAGACGCGCG CnTGCCCAAA CACGACGTAC ACCGTCGAGA   840

TTCCACCGGT TGTCGAAACG GCCCTAAAAA GGCCCGAGTC TGAGGGGAGn TCACGGTCTC   900

nGGAnGCCTT ACAGGsGTCA CGCTGCACCT TTTTGCGCTC ACTCTTGTCC TGCGCAATGG   960

CGGTATTCTT gCGATCTAAG CGTAATTCCT CACGGGGAAA TTCCTTTTTC TCAATGCGCA   1020

GCGGCGCACG CCGCACATAT GGCTTCCCTC CCGGCACACG CACCTGCGCT TCAAGCATCC   1080

AATCCCCCGG TTCCCAGAAA ATCGATATCC CCAGCAAGGC AACGTGCGTC ACATCCTGAG   1140

AACCTGCACG CGAGACGCCA GCGGTTGCCG CGTCCGTCCC TAACTGAGCA ATACCCTTTG   1200

GGGGAAGCGC AAAAGCGCGC ACCGTCTTTG CTTCTTTACC CGCAGGGGTA CGCAGCACCA   1260

GATGTACCTC AGTATGCGCC TTGTCCTTTT CTTGCAATGC CACTAAAGAA AAAGTGGCCA   1320

TCGCACACGC ACCTTGGGAT ACCTGACGCG GGAACTGCAT AGCGATACGC TCGAAATGTG   1380

CAGGAACCAC CTGACGCTCC GGCGGcGGCA CCGCAGcCGA ATGAAGCAGG AAGGACACCG   1440

CGCTGACAAA GACACCAGAG AACAAGAGTA cTTCGCACAG ACCACGCACC CAACGAACAC   1500

TTCTTTTCAC CGACGGTGAC TGCACGCCCT GCGTCTGCAC TGCCCTTTTA GCGTTCACCC   1560

CCGGTGCGCG TCCTACTCTC TCTGCACCCA TCACTCACTC CTCCACAAAT CTTGAATGAC   1620

CAGCTGCGGT GTGCACGTTC CCTGAAACGT GTTACGCGTC ACTTGAAAAA CCGCGTCAAC   1680

TACATCCCCC ACTGCAAACT CCTGTGCCAA CTTTTCCCCT GCTCCCCAGT AAATTGCGGG   1740

CCATTTATGC ACCTGTGCAT CCAAGGTCAA TTTTACGTGC ACACGTTCTG TACGCCCAAA   1800

AAGCGATGCA GAAAAAATTT TCAATCTCTT CGCCAAAAAG CACAACGGGG GATTGCCTTC   1860

TCCGTACGGC TCAAAGCGAT CGACAAGGGT CAAAAGCCCC CGCGTCATCT GCGTAGCATg   1920

```
cCAGTTCTGC ATCAAATTCT CCACACTCTT GCGCGCTTTC ATCAGCAAAC TCAATGGTTG      1980

CCGCATACAG TTCCATACGG TGCAATAGCT GGGGAATTCG CTCAGAGGGA ATTGAAAAAC      2040

CCGCCGCAAA TGCATGCCCC CCATAGTCAG AGAACAAGTC TGCAAGGGGA TCTAAGAGCG      2100

AAAATAGGTG ATATCCCCGC GCCGAACGCA ACGATCCTAC CGCGTGCCCG TCTGCCATTA      2160

TACAAATGAT CACACAAGGC ACGCGCAGCC TnCGcTCAAA CAGTTTGCAA GAATCCCCGT      2220

AACGCCCCGA TGAATCTTAT CGCTACAAAC CACTGCCAGG CGGTTGCTGT ACGTCTCAAG      2280

ACTTGcACGC GCAAGAGGCT CAACAAGTGC ACGAGCACTC CTTCCTAACT TTTTTCGCTG      2340

TTCGTTCAAT TGCACCATTT TTCGTGCCTG cAGCGCGCGC TGCGAAgTTT CGCGCATTAA      2400

AAACAGTTCC ACTGCACGGT GCGGACACCC TAACCGCCCC GTTGCATTGA TAAGCGGCAC      2460

AATACTCCAC CCTAnCTCTA CGGTTCCTAA CTTCTTCCCC ATGAGACGCT GTATCGcAAA      2520

CAAcTcACGC AAACCCACAC GTGGACGGCC TCATTCATCG CCTGCAGACC GTAGCGGACC      2580

AT                                                                    2582
```

(2) INFORMATION FOR SEQ ID NO: 60:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 5504 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

```
CAAACAGATC AGCGGGAGAA ACGTCTGGCA ATTnTAGCGC GCGCGCATCG GCAGGAAGAG        60

ATGGCGGCAG CCCCGGTTTA TCTGCCACAA GCACACAGCG CACCAAGCCC CCGCGACGAG       120

CAAGACAAAG CGTCTGCAAC TCATCGGGCA AGCAGCGGTA TGAACGCTCA GAAAGCTCCA       180

CGGGGATAAG ATGTATCCCC CGCTGTGCGA GTGCCCTCAC ATCTGCCGGA TCGAAATCCC       240

CCCACGGCTC ACAGCGTTCC AGATGCGCAA TGCACTGCGC AATGCGCTGA GCAAGTTCCA       300

CGCGGTCTGA ATGAGTACGT ACGATCTGTT CCACGGCCTC TGCTGCTTCT ACCACCTGAC       360

CTGCAACGCG ACACTCCTCT CCTCGGGTAA CATTCTTTGT CTGAGCGTCA GTGACGAGCG       420

CAATGGCCTG CACGCACCGA GCGTCCAACG CGTgCAACTC TGCAAGTTGC TCACTTGCAC       480

ACTCCCGCAA CTGCACATGC ACAGCACCAA AGGAACGCAA CGCCTGcAGC GAACGCTCTT       540

GCTCAGAACC GAGCACCAGA AGCGTTACCT TTTTCATAGG AACTATCACC GTGTATCCTC       600

TTGTTCCATC CGACTCACGT CTACCAGATT CTTCTTAGAC ATCTTCCCCC GCACTACTGC       660
```

```
AGCAACCTGC TGATCACCGA GGTACACCGT TATTTtCCGT ATGCATGCCC GCGTCTCAGG    720

AATCTTAACT TTTTCAAAGA GGTTAACACG CTGCGTTGTA GTCCGCAACT CTGcACCCAG    780

GAGAAGCGCC TGTTCGTCGA GAACATGCGC CTCCAAGTCT AAGCTTAGCA CTTCCTGCAT    840

TTTGCGCACT GCAGTATCCA CCCACAGAGG AACACGATAT AAGTCATAGG GAGGACAAGC    900

AAAGTGCACT TCTAAAAAGC AGGGAATACG CACACCTGCA ATGCTAGCAT ACGTTTTCTT    960

TACCTCTTGC ACGCGGAGCA AACGCGCGTC GAACACACCG CTTTCAGAAA AAACTGCAAC   1020

CCACTGCTGA ACATCCTGAC GCAGGGCATC TGCACGGGAA CGTACTTCAG AAGCGCGCGC   1080

CTCAACGGCA CGGATCTCAG CATACAACTG CTGCTTTTTA AGCTGAAGCG TAGGGAGAAA   1140

ACGGCGAAAC GTCTTGAGCG TCTCTTTTTG ACGTTTCAGT TCATTTTTGG TTAAGCGCAC   1200

CGCCAtCGGT CACCACCCTA CGCAGGCCAA TACGTGTTAA TCAAATCAGA GCGAATCCCC   1260

GTCTCCTCTG GGGTGAAACA CCGGCCCAGA ATTTTCCACC CCGTATCGAA CGCCTCTTCA   1320

AGCGGAATAT TCACCGAAAG ATCCATGAGc TGCGCTTCAA ACAGCCCACC GTATGTGAGC   1380

AGTTTCTCAT CCCACTCGCT CATGGCAAAA CCCATAGATC TTTTCTCAAG CGCATCACGA   1440

TAGGCGGCAT ACAACTTAAT CATATTATCC ATAAGCGCGC GATGATCTGC ACGCGTACGC   1500

CCGTTTACGT TCTGCTTAAG ACGGGATAGA CTCCCGAAAG gTTCAATGCG CCCGTTCTtc   1560

AGATAAAACT GaCCCTCAGT AATGTACCCC GTGTTATCAG GAACCGGATG CGTAACATCA   1620

TCCCCTGGCA TGGTGGTAAC GGCAAGGATA GTCACTGACC CTGcATCATC AAAATCGACC   1680

GCCTTTTCAT AGCGCGACGC AAGCTGGCTG TACAAGTCAc CCGGATACCC ACGATTCGAG   1740

GGAACTTGTT CCTGAaTAaT CGCAaTTTCC TTCATAGCAT CAGCAAAaTT AGTCATGTCG   1800

GTTAAGAGCA CCAACACATC CCTACCCTTC AAGGCAAACT GCTCGGCAAC TGCAAGaCAC   1860

ATATCAGGGA CCATCAAACA TTCTACGGTA GGATCTGAGG CAGTGTGCAC GAACAGGACT   1920

GCCCTACTCA ACGCTCCTGC CTCTTCCAAT GCACTTTTAA AATACAGGTA ATCGTCATGC   1980

TTCAGCCCCA TACCCCCGAG GACGATGACA TCAACCTCCG CTTGCATTGC AATACGGGCC   2040

AGCAGTTCGT TGTACGGTTC CCCTGAGCTA GAAAAAATAG GCAACTTCTG AGAAACAACC   2100

AGCGTATTAA ACACATCAAT CATGGGAATA CCCGTGCGAA TCATACGCCG CGCGATAACC   2160

CTCTTTGCCG GATTAACCGA AGGACCGCCA ATTTCCACCC TCCCTTCCTT TAAGGCCGGA   2220

CCACCGTCTC GGGGAACGCC AGAGCCATTA AAAATTCTCC CCAATAAATA ATCTGAGAAA   2280

CTCACGAGCA TACCCCTCCC CAGAAAGCGC ACCTCGCTCC CGGTGGAAAT ACCCCGGCCT   2340

CCCGCAAACA CCTGcAGGGA AACTACATCC CCTTCAAGCT TATTCACCTC AGCAAGCGAA   2400
```

```
TCGCCAAACG CCGTTTTTAC CCGGGCCAAT TCCCCGTAAT GCACCCCCTT TGCCCGCACC    2460

GTGATGACAG AACCGTTGAT CGACTCAATC TTCTCGTACA CCTTGTACAT CGTCTACTCC    2520

ATCCCCCGTA TAATTCCCTC TGcTTCGCTG TCGATTTTCG TCGATTCTCC CTGGAGAAAG    2580

GCACGTATCT CCTTTTCTTT CTCCACAAAC GCCTCAGAAT TCCAGGCGCA ACAGTTGTAA    2640

TCGATAAACA TATGCCCAAG CTTGCTGAAG TATGCCCGCG CGTCATCTTT TGATTCAAAC    2700

GCTAAAACAC TGCCAAGAAC CCGCATGACG ATGGCATAGC AGTGCTTTTG ACGTGCAACA    2760

GGTACTGCAC TATCGACTGT GTCAAAAGAA TTCTGCTGCA GATACACCGA ATCAAGAAAC    2820

GAGCCTTTCA GATATACGAG AAAGTCCTCC ATACTTGTGC CCTCTTCGCC GACGACCCTC    2880

ATCATCTGCT CCACCTCTGC CCCACGGCGC AGAAAAGAGC GACCGTACGC AACAGCCCGC    2940

GCGTCAAGCA CACTTGGATA CTTAGACCAT GAATCAAGCG GATGCACCGC AGgATACCTG    3000

CGCGCGTCAG AGCgyTnCTC GAGAAAGTCC GTGAAAgCCC CAACCACTTT CAATGTAGCC    3060

TGCGTTACCG GTTCTTCGAA ATTACCACCT GcCGGAGAAA CCGTCCCTCC AATAGTTACC    3120

GATCCTTTCT CTCCACTCCT CAGCCGGACC ACACCAGCCC GCTCATAAAA GGCTGCGATA    3180

CAAGACTCCA GGTACGCAGG AAAGGCCTCC TCCCCCGGAA TCTCTTCCAA ACGCCCAGAC    3240

ATTTCACGCA GGGCCTGTGC CCAACGGCTC GTAGAtCCGC CAGCAAAAGA ACATCCAACC    3300

CCATCTGACG GTAATATTCT GCAAGCGTCA CTCCCGTGTA CACTGAAGCC TCACGAGAGG    3360

CAACGGGCAT AGAAGAAGTG TTGCACACTA TAACCGTCCG CTCCATAAGC GACCGACCAG    3420

TGCGAGGATC CGTAAGATCA GGAAACTCCC GCAGgTtCTC AACCACCTCC CCTGCACGCT    3480

CCCCACACGC AGCAATCACT ACCACGTCCA CATCCGCATT GCGACTGGTA GAATGCTGCA    3540

GCACCGTCTT TCCCGCACCA AAGGGACCGG GAATACAGTA CGTCCCCCCC TTGGCCACCG    3600

GGAAAAAGGT ATCTATCGTC CTAATGCTCG TTACCAATGG CTCAGTCGGT TTCAAAcGCT    3660

CTGcGTAACA ATGGACGGGT CGCTTCACTG GCCAACGAAA TGCCATGGTC AGTTCGTGCT    3720

CATGTCCCTG CGCGTCACGG ACCCGCGCAA TCACATCGTG CACGCGGTAC GTCCCTGCAG    3780

TCTGAATGAA GACAACCTCA TAGGAATCCC CCATATGAAA GGGAACCATA ATGCGGTGTT    3840

TGAGCGCACC CTCTGGGGTA TACCCGAGCA CGTCCCCACG CACCACGCGC TCACCCACTG    3900

AAACATGCGG GGTAAACATC CATTCACTTG TCCGAGAGAG GGCGGGCAAA TACACCCCGC    3960

GCTCCAAGAA ATACCCAACC TTTTCTGCAA GCAGCGGCAA CGGATTCTGT AAACCGTCGT    4020

ACACCTGACC GAGCAAACCA GGACCTAGCT CAACAGACAG CAAATCGCCT GTAAACTCAA    4080

CACGGTCCCC AACGGAAACC CCTCTTGTGA TCTCAAACAC TTGCAACTGT GCCTCACGAC    4140
```

```
CACGAACACG AATAATCTCC GCTTTCAAAC GCGCATTTCC AACATGCACG TATCCGACCT    4200

CGTTGAGCGA AACGACACCC TCGAACGTAA CGCTCACCAT ATTGCCGTTG ACCGCAGACA    4260

CGATACCcTT CGTTTGCGTC ATGTATATGC TCCAAGAATT GTATAATAAA GAGCTGCGTA    4320

TGCTTTGGAC CCCGCCTGCA TTGTGAAACG CGACCGATGC GTAAGCAACA TCAGCTGCAG    4380

CCCGTACAAA AACACTGCCT CTGAGGAAAA CGGGTCAAGC GGCCTACACC CCTCGACAAA    4440

GGAAAAGCGC GCGTCGTTCA AAAAATACTC CGCTTCGAGC GGATCGTCCA AAGAGACCGC    4500

AACACGAGCC GCACGAGCCA CCGACTCCTG CGCCACAGGA CACTGCCTTA CTTCAACAGG    4560

AGTATCCCAC CGCAGACGGT cCgcGCGCTC ACgcgCAAGc GCACAGCGTA ACGCGTACTC    4620

AAACTCGCCC CATCTATCTA AAACACGCGA TCCCGTGGAG TCACGCGCCG GCACGGGACA    4680

CAACGAGATA TTTCCAAGCA CCGCAGCATC CTGCCGACCC AAGAAACGTA GCGCACAATC    4740

CAAAAAATCC TGATAACGCA AAGGAGGCAC CGCGCCGCAT AGAAGAGATG GTAGCTGCGT    4800

TATAAGGTAG CAATAGGAAG ACATCACAGC TCcTGCGCAG CAGmCTGAGC ACCTCGGCAA    4860

CACGCGCAGA AACATACGAA GAGAACAACT GAGCAACgCG GCGGCGGAAA AAtcATAGTa    4920

CGACCCGCCC TCGGCAGGGA CTATCCTAAA CCCTGCCGTA AGGCAATCGT CAGACCTCAA    4980

CTCTACCCCT GCCGAAAGCT GCTCCTGCAA CGCAsCACAA AACACCCCCT CAAGCGTCCG    5040

AAGATCAGCA GGAGAGAGGA TGAGCTCTAG CTTATCACCC TCCGCTTGAA CCCAGGCAGA    5100

AACGACACGA GGAATAAGtc ACGCAAAACA CCCGCATCGT aGtTGCGCCG TCTCCATCGA    5160

AATAATAGCC CGAAGAGAGC GAGTCACCGA ATCTTGAAAG GATAATAAAA CGTTGCGACT    5220

CGCCTGCGAC AGGCCGCAAG AGACGACGAC TCGATCCGCT CTGCCTCCTC ACGCGCAGCG    5280

GAACAATCCG CTCTGCCTCC TCACGCGCAG CGGAACAATC CGCTCTGCCT CCTCACGCGA    5340

CTCACCAAGC AAACGAGACG CCTGCTCCTC GGAAGAGGnC AAnCGCTTCG CGCTTAATTC    5400

GGTCATCAGA TCTTGCAGTT GAATCTCCAC TTATAGTTCT CCTCGCAGCC CTCCGAGTAT    5460

ACTAAAAAGT CCCCACCGGG AnAAGGCATA ACACAnTTCG ACCA                     5504
```

(2) INFORMATION FOR SEQ ID NO: 61:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 8467 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

```
TTGTATTAAC CCATTGCCTT ATCCTTTTTC ACCCAGCGCC AGTTCACGAG ATGCATTACG    60

TTCCTCCCTT GGAAAACGGA GAATGACTTC CGTTATATCC GCCCGTTCTC TAGGGTGGAG   120

ACAAATCCAT AAAAGTAACG CCTCTTTTTT ACTCCCCCAT ACCTCATCAC CGCATACAAA   180

GCAAAACAAA ATCACTGAGG TTAAACATAC CCACCGTGTT ACGCTGTACG CGAATCCACA   240

GATCGCATAA CCCTCACCGT TTTCTCCGAT AAAGAATCTG CATCACCACA AACAGCATTC   300

CTATAGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA   360

TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA   420

CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG AAGaGCTAAC CGCTTTTTTG   480

CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC   540

ATACCAAACG ACGAGCGTGA CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA   600

CTATTAACTG GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG   660

GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGTATTTGTT CCTTACCAGG   720

ATGCGTACTC CCCTTCGTAC AGCGCCGCTT CTCTTGCTGC TCCTATGCGC ACTTCCCCGG   780

GCGTTGTGTT GCTCTCTAGT GCAcTGCGCG GGGTACCATT CGATGTACCG ACCCCATACG   840

TTTCCCGTCG GGCGAACACT ATCGACGCTG CCACCTTTGA AGACGCTCAT GTACCTGCAT   900

TATTTCCCGC GCTCTTTGCG CTTTGCAGGC ACGCGCCCaC ATTCGTGTAC GCAGAAAGTG   960

CCCATGAGGT GATGCTCAGC CGTTTTCTGC AACAACAGCC ACATGCATGC GCCGGTGTCT  1020

TTTTTGTCCT TCCTGACTCT GCAGCGCGCG GACCACACCA TGCTCCTGCC GTGCAGGGCG  1080

CACCTCCCCC CGTCGACACA GCGGGCGTTG CGTCTGCTGT CCGTGGCGCC AGCCGGACAC  1140

TACCAGCTGT GTATCGACAG TATGTTCACG CAGCAGAGGC AGCGTGGGCA GAGCTCGCAT  1200

CCACCGATAT ACTGGCCGCT TACttGCAGG GCTCCCTCGG GACCGCCACA GAACGCGCCT  1260

TCAGGCACGT GctACGCAGG TAGACCAGTG GATACGCGCC CAGCTGCATC TATCAGAACC  1320

TGTCCTTCCG CACGCGCAAG CGCTATCTCA TCACACTGTC CATGCGGGAG GGACCTATGA  1380

CCGCACGTGA GTTAGATGCG TACTTTCGTA GTTTTTTGAA CTTTGGACCG TTCGTCTCCT  1440

GTGATGTCGC TCTCAACGGC CTGCAGGTAG CAAATAGCGG TGCCCCCGTG CACAAGGTTG  1500

CCTTTGCAGT GGATGCGTGT GCACAGTCTA TCGACGCAGC CGCCCGCGCC GGTGCACGCA  1560

TGCTCTTTGT CCATCACGGT CTTTTTTGGG GACGCATAGA GCCGCTTACC GGTATGCAAT  1620

ACCGACGCGT ACAGGCGCTC CTGACGCACG ACATAGCGCT GTACGCAtGC ACCTACCACT  1680

CGATGCACAC CCGCAGTACG GTAACAATGC GGGCCTTGcT GCGCGAGTCG GTCTTAGGCA  1740
```

AGGTGGTCCT TTCGGTTTTA TCCGTGGAAC TGCCGTAGaC TCTGGGGGAC GGTGGCAGAA 1800

AACACCACCC CCTCTCAGGA GGCAATGCAG CAGCATGCAG CGTGCACAGC ACCCGATACC 1860

CACCGCGTGA CGCATGCGAA TGCAATATCG CCGAGTGCCG GGCTATCTCT CCAACAAGTA 1920

GTACATCGCC TCTTCCCCGC AGAAGAGCAA CCCGTGCGCC TGTTACCGTT TGGGAAACAG 1980

CGTATCGAGC GCGTGGGTAT AcTGTCGGGC AAAGCAGGCA CGTACCTTGC GGAGGCTATC 2040

GCGTTAGATC TGGACCTGTT TATTACCGGG GAGATTGAAC ATTCTTGCTA TCACACCGCG 2100

CGCGAGCACT CTATCTCGGT AATCGCAGGG GGACACTACC AAACAGAAAC CGTAGGtTGC 2160

AGCTGGTGGC GCGCAAcTGC AACGGGATAC AGGCATAGAA ACGCTTTTTC TAGACATTCC 2220

CACGGGGATG TGATACGCTC GCGCCCGTTA AGGGTGGATA CAATGAAACT CACACGGATA 2280

CAGAAAGAAA AGTGGATCCC GCTTTTTGCC GCTGGATTAG TTGTTGTTCT GGATCAGTGC 2340

GCTAAATTGT TGGTGGGTGC TTATGTGCCT ACAAACACCT CGGGCGTTCG CGTGCTCGGT 2400

GATTTCGTGA GAATTGTTCA CGTGTACAAT GTTGGCGCCg CTTTCAGCAT TGGCCATCAG 2460

CTAAATCAGG TTCTGCGTAC GCTCGTGCTC GGTATCGTGC CGCTAATCAT TATGTTCCTT 2520

ATTGTTTTCT CCTATTTTCG CACTGACGCC TTCTGTCCTG TTCAGCGCTG GGCCGTGTCA 2580

GGGATTATCG GGGGAGGGAT AGGGAACTTA ATCGATCGCT TCCTGAGGCC AAACGGGGTG 2640

CTCGACTTTA TCGACGTAAA GTTCTTTGGC ATCTTTGGCT TTGAGCGCTG GCCCGCTTTT 2700

AACATTGCAG ATGCGGTCAT CATGACCTGT GGTTTGCTCT TGATCATTTC GTTCATAAAA 2760

CAAGAAAAAG AGATCAGCTC CCAACCCTCC TGCAATGAGA CGGGGGGCGT TTTTCGCACG 2820

TAGAGCTGGG CCGTGCGCGC ATGTCCGCGT CGGCCGTTCT AGTTCGCGTG CCCCTGTGCC 2880

CGCAATGGTT GCTTTGTTCT CCGCAAATAC CGCGCGTGTG TGCCGCGCGT TGCgcTtCCG 2940

GCGTACCAGG GCGGTACgcG CGAGsgcCTC ACAGCACTCA GGATATTAGC CCATGCAGAT 3000

CTTCGATACT CACGCCCACA TCGGTCTTAT TCACCCAGAT CCCGTAGAGC GGCTGCGgGT 3060

AGTACAAGAG GCACGACGAG CTTCTGTCAC CCGCATCATG AGTATTTGCA ACAGCCTTCA 3120

TGACTTTGCC GCCGTATACG AGACGCTCCA GTTCTCACCC TCTGTCTATC ACGCCGTAGT 3180

GTCTCCCCTT CTGAGGTCAT GGCCCCGGGG AAGGATTGGA TAGATACTAT TCAAAAAAGC 3240

CTACAACTCC CTCAGGTAGT TGCCTTAGGC GAGACCGGAT TGGACTACTG TAAAAAGTAC 3300

GGTGATAAAC GCTCCCAGAT TGGGCTTTTT ATCACTCAAT TGGATATTGC TTCAAAGGCA 3360

AAAAAACCAG TTATCATCCA CAACCGTGGT GCGGGCCAGG ATATCCTGGA CATCCTCAGC 3420

GAGCGCATTC CCGACCAAGG CGGTGTGTTC CACTGTTATT CTGAGGACGC AGAGTACGCA 3480

```
CGTATGGCGC TGGATTTACC TGTGTACTTT TCTTTCGCGG GGAATTTAAC TTACCGGAAT    3540

GCACGAAATC TCCATGAGAC TGTATTGGCC CTCCCGCTTG ACCGAATTCT AGTGGAATCC    3600

GAAAGCCCGT TTATGTCCCC CGCCACGTAC CGCAACAAGC GCAACCGACC GGCGCACACA    3660

GTTGAAACCG TGGAGTTCAT GGCTGAGCTC CTTGATATGG ACATGCTTGA GCTTGCCGAC    3720

CAGCTGTGGA AAAACAGCTG TGCGTGTTTT CACCTTCCTG AGTGAGCAGC AGATGCAACA    3780

ACACGCCTTA TATCATCCGG TTTCTATTGG CCCGTTGTCT CTCAAGGGGA ATGTGTTTTT    3840

TGCTCCCGTT GCAGGCTATT CTGACAGTGC GTTTCGTTCA ATTGCCATTG AATGGGAAGC    3900

AAGCTTCACC TACACCGAAA TGGTTTCGTC TGAGGCGATG GTGCGCGATT CACTCAATAC    3960

CAAACGTTTG ATTCGGCGCG CGTCAAATGA GACGCATTAC GCTATCCAGA TTTTTGGTTC    4020

TAATCCTGCA GTAATGGCAG AGACGGCAAA ACTAATCGTC GATAGCGCGC AGCCGTCCTG    4080

TATCGACATC AACGCGGGAT GTCCTATGCC TAAAATCACT AAAACAGGAG CCGGAGCCGC    4140

ACTCACCCGA GAACCGACGC GCCTCTATGA AGTGGTAAAG GCGGTCGCCG ATGCTGTGTa    4200

CgcGCAAGAC GCGCGTATCC CAGTGACAGT AAAAAATTCGT GCTGGGTGGG AAGAGGCACA    4260

CCTGACATGG AAGGAAnsTG CGCGTGCGGC AGTAGACGCA GGAGCACAAG CGCTTGCGTT    4320

GCACCCgCGC ACCTGCGCGC AGTGTTACGC GGGAGAGGCA AACTGGGACA TAATCGCAGA    4380

CCTCGTGCAG TGCGCGCGTG GGTGGGGAGA GGTTCCCGTG TTCGGCTCAG GGGATCTGCA    4440

TGCGCCTGAA GACGCACGGG CAATGTTAGA ACACACCGCA TGCGCGGGGG TTATGTTTGC    4500

CCGCGGTGCT ATGGGCAACC CGTTTATTTT CAGACAAACC CGTCAGCTTT TAACTGAAGG    4560

ATACTACACG CCCGTGACGT TTGAGCAAAA GcTACGCGCA GCCTGGCGCG AGCTTCACCT    4620

TCTGGCACAA GACGTGGGAG AAAGCTCAGC CTGCAAGCAG ATGCGCAAGC GTTTTGTTTC    4680

GTATGCAAAG GGTGAGCGGG GTAAAACGCA ATGGTGTCAG CGCGCGGTGC ATGCGTCTTC    4740

CTTCGCAGAC TTTGCAGCAG TCATTCGTGA CGCGTGTCCA TGTATTGGTT TATAAGTTGC    4800

ACGGCTTTTC AAACCGCGTG AAAAACGTAC GCTTCCGGCG TACCCCAACT TACTTTGTCC    4860

TACAGGACGC GCAGnTCCCT CGATAGAAAG CGTGACTATA TCTGTCCTGC GTGCAACTTG    4920

TACAAAGCCG GTTCCTGTGC CAGAATCGTG CATCCGCGGT GGCAGGGAAT CCTGGTGAAA    4980

GAATGTGTTT CTAAAGAAAA GCGAATCTAT GACCTCAAGC AGCTCCTAGA GATTTCTAAG    5040

AGTTTGAATT CTCTCCTTGA GTTTACTCAC CTGGTAGAAG CCATCCTCTA CGTCGCGATG    5100

GCCCAGACCA AGACGCTGGG GGCAGCGCTT TTCACCAAGA AAAACGCCGG TATGAAAAAA    5160

TTGTCTTTGA GCCGCAaTGT GTGCGGCTTT GACGTTTCCC ACCATGCACA GCTGATAATC    5220
```

```
TCGGAAGAGG ACCCTATTCT CAGACTTCTG GACGAAAAGG CCTGTTGTCT TTCTCCCgAA   5280

gAgGTACAGA GCGCGCTCGC CCCCTCAAAG AGCGTACGTT CGCTCCyTGA CTTGCAACCT   5340

TCGCTCTTTG TTCCACTAAG AGCAAAGGAC CACCTTGTTG GTCTTATCCT TTTAGGCAAG   5400

AAAAtCAACG TACACGAAGC CTACACTCCC TACGATCAGA GCATCATCAT GGATATTGCA   5460

CAGCTTGCTG CTATTGCCAT CAACAATGCG TTACTGCTTG AGCAAGCTAC CACTGACATG   5520

ATGACCCAGA TGAAGCTCAA ACACTACTTC TTTGCCATGC TCACCGCGAr CTCGATACAC   5580

TCAGTACACA AGAGACCGTA TCTGTTCTCA TGCTTGATAT CGACTTTTTC AAACAGATCA   5640

ACGACACGCA CGGTCATCTG TGTGGCGATC TAGTTCTCCA ACATGTGGCA GAAATTATTC   5700

GATCCTGCAC CCGTCCATGC GACATCGCCT CTCGCTATGG GGGAGAAGAA TTTATGCTCA   5760

TGCTATCCAA CAACTCGTCT CGGGaAGctG CGCACGTTGC AGAAmgCATT CGCGTGGCAA   5820

CCGAGCAATT GACCATCCCC TACCATGAGG TATCAATTCG AGTCACTGTT TCTGCAGGCG   5880

TCGCAGAATA CCTTCCTAAC CAAGAATCCG CCGAAACACT GATAAAGCGT GCAGACAGTG   5940

CGCTGTATCA AGCCAAACAA AATGGCAGAA ACAAAGTCGT CATCTCAGAG AAAAACATGT   6000

GCTCATCTCA GGAATAAACC GATACTGGCG GCATGAGTGT GATCAGGAAG CCCTTCAGGT   6060

ACTCGTACAC CAATGTGACC CTTTCCCTTG TGCTCGCGAA TGGGGCGGTG TTTGTGATCA   6120

CGTCGTTGGT TGAATCACTG GGTATATATC TGGCGCTCGT GCCAGGACTC GTACGTTACC   6180

ACCGTATGTA TTGGCAAATA TTCACCTATC AGTTCGTACA CAGCGGCGTG TGGCACTTGC   6240

TTTTTAACAT GCTAGGACTA GTGTTTTTCG GGCAGACGAT AGAAAAGAAG ATGGGATCTT   6300

CTGAAATGCT GTTGTTTTAT TTGCTTGTCG GTACACTCTG TGGTGCGGGT GCGTGCGCGG   6360

CATATCTGTG TGTCGGTCGG TTGAACGTAC TGCTGTTGGG GGCGTCGGGC TCCATCTTCG   6420

CAATACTTTT TTTATTTTCG GTTATGTTCC CCCACTGCGC TCATTTATCT ATGGGGTGTT   6480

ATTCCTATCC CCGCTCCTCT GCTCATTGTA GGATACATTT TGTTTGAAAT TTTTGATCTA   6540

TTTTTCTCTC GTGATAATGT TTCTCATCTT ACCCACTTGC TCGGTGTCCT TTTTGCGTGG   6600

GGATATATCC GTATCCGGTT TGGCATCAAA CCATTGAAAG TGTGGAGCAT TGTCCCGTAA   6660

CAGTCGAGGC AGTGGGAGAT ATGTCTTCGT CGTGCTAGCC TGCGTATTTG GTTATACGCG   6720

CGCCGTGCAC GCTGAGGTTT ATACGGACCC CAGCACATCG GGACATGTCA CGATTTCTAT   6780

TCCCATATGG GCTTyTGTCG AGCCCCAGCC GGGTGTCATG ACCCAGcAGs GGAGTCCCCG   6840

AGGACTCCGC CTCnCCAGAC CTTGCGAGAA TTAGGGGCGT TCGTATTAGG CGGTGCTGTG   6900

TATGGGTGGC GGTTCTCTTA TACGCCaAAA GAAAAGAAGC GCGCCGTCAT GGAGCACTTT   6960
```

```
ACCCTCACTC CCATTTTCCC CCTACCGCCC GATAGTCCTC AGATAAGTCT GCGTCACGTA   7020

CGGACGCCGT ACCCCTACAT CCAtGCCGTG CAGAGTACTC ATTAGACGCC AGGCACGCGA   7080

CACACATGAG ACAGAGCAGA AACCTAACGT ACCAACGTGC GCAGGGCAGA GGAAGAGGAG   7140

AACGGAAAGA GGAACTAAAG GGAGTATATC ATGCATATCA CCGCGCGATT GTAGACGCAC   7200

TACGGAAAAC GGTTAGAAAG ACACAGAAAA ACAAGCCAAA AGAAGTAGAA GGAATGCTAT   7260

ACGTTAAAGA CAATCCCCGC CTCTTTGTAG AGGCGGGGGA ATTTGTCGCA GAGCTCTCAC   7320

TCAGTGTCCA CTTCACAAAG ATAACGCCCT ATAGCGTATA CTAGTAGCAC GCACCGAGTC   7380

CTGACCGCTA CCCGCGTGCG AGCAGACGGT TCACCCGCTT CACAAAATCA ACCGACGAAC   7440

CTACGTCCAT GCCTTCAATG AGCAAGGCTT GATCCAGAAG AACAAACGCA AGATCTTCCA   7500

CAAACGCCTC ATCCGTACTT TCTTTTAGTT TTTGTACCAG CGTATGACTT GCGTTAATTT   7560

CTAAAATTGG CTTTATCTTT GATTTATGCG TTTGTCCCGT GGCGCGCATC AAGCGCTCCA   7620

TCTGCACCGT GGGATCATTC TCATCGATAA CAATGcAAGA CACCGAGTCA GAAAGCCGTT   7680

TTGAAAGACG AACTTCCTTC ACCGAATCAG ACAGTATGTG CGTCAACCTT TCTAGTAGCG   7740

GCTTAAAACC CTGTTCCCTC TGCGCGGCGG CGTCTGTTTC TTCGTTGGGA CGCAACTCCT   7800

CCTCTGAACC TAAACGATTA ATTGCCCTTA ACTCCCACTC CTTGTATTTC GAAACAGAGG   7860

GCATCACGAT ACCATCTATG TCGTCTGACA TAACGAGCAC TTCAAAACCC TGcAAACGAT   7920

AAGACTCTGC ATGGGGAGAC TGACGCAGCA CACGATCGTC GTTTCCCGCA ATGTAGTATA   7980

TCGCCTTTTG ATCCGGTTTC ATGCGAGAAA CGTATTCGGC GAAcTCGTCC ATCCGTCTTC   8040

TGGAACAGAC TCACTTAGAG TCCTGAAACG AACAAGTTCC AGCAGCTGCT CACGGTGCTC   8100

GTAGTCGCTG TATAAACCCT CCTTCAAGGG ACGATTATAC TGCGTGATAA ACTCATCGTA   8160

CTTTTTCCCG TCACACTCCG CGAGTCTCTT AAATTCCCCG AGCAACTTTT TCACCGAAGC   8220

CGACTTGATT GCTGCAAGGA CTCTATTTTG TTGCAGAATC TCACGGCTTA CATTCAGGGG   8280

CAGATCTTCG CTGTCTATTA CACCGCGGAC AAAACGCAGA TACACTGGCA ACAGTTCCTT   8340

CTCGTCATCA GTGGATGAAA ACGCGCTTAA CGAATAGCTT TACCCCCGGC TTATAATCTG   8400

ACGTGAAAAA GGTCAAAAGn GCGCTTTTTG CCGGGCAAAT AAAAGAGCGT nGACGTACTC   8460

CTGTGTA                                                            8467
```

(2) INFORMATION FOR SEQ ID NO: 62:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4354 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

```
CTCTTCAATA ATGTCTTCCA TGCACGCAAT ACCCGAAACG CCGCCGTACT CGTCCACCGC      60

GATCGCAATG TGCACGTGCC TGCGCTTAAA CTCTCGCAGA AGACTGTCAA TTCGTTTGGA     120

CTCGGGGACA AAGAAGGGTT ACGCAGCAGT CTTTCTAACC GCACCTCCTG TGGCCTTCCA     180

AACAGCTTTA TTAAATCTTT GACGTACAGC ACACCCACCA CATTATCAAT AGTTTGTTCG     240

TAGACAGGAA AGCGTGAGTG TCCACTCTCG GTTACCTTTT CAACGAGTGT TTCACCGCTC     300

ATAGAAAGCT CAAGAAAATC CACGTCAATA CGCGGTATCA TCACCTCGCG CACCGAAGTG     360

tCAGAAAGAT CCACTATACC GCGGATCATA TCCTGCTTTT CTTCATTCAG CGGTTGCTGA     420

AAAATATGGG TAACAGCGTG CCTGCGCCTC AACCAGTCTA TGACTCCCAT GGTATACCCG     480

ATGATAGCAC CCGACAGTGt GCGCCAGTAT GCGCTCCTGC AAACGCAACA TCTCTTGACC     540

GGGAGAATTG TCCTGGTGAT CCATACCGCT CAGATGCAAA ATGCCGTGGA TGAGCACCCG     600

TTTAAATTCC TCGTGCGCGG CAACGTGAAA ACGTTCACTG TTTTCACGCA CACTTTCAAG     660

ACTGATGATA ATATCACCAG CAAGAAAAAA ACGCGTCCCT GCGTCATCGC AATACTCACC     720

ATCGTTCTCA AAAGACAGCA CGTCAGtGGG AGAATCAATA CCACGGTAAT CGTAATTTAG     780

CCGGCGAATA AACGCATCAG TGCAGCAGAC AATGGAAAGA TCCCAGTGGG AAATAGCCTG     840

GGAATCGAGC ACCGCACACA CAAACGGCGC AACTTGACCA ATCCAAGGAG GCGGACAAAA     900

GCcTTCGCAG GAAACAGAAA CTTTATTCAC CTCGGACATA AAGATTACTC CTTATACGAT     960

CCTTGGGCTA CGGACACGAG CTGCTGCTGA TCAGAATCTC TTTGGTGAGG ATACTCTATG    1020

CGGGAATGGT AGTATCCTGC CAGTATTCTC ACAAAACACT CCTTGACGAC CTGCACATCC    1080

CGAAACGTTA AATCAGAATT GTCAAGCTGa TGTGTTTCTA TCTTCTGTTG CACAACCTTA    1140

TCGATAAATT TCCCTAGGCG GGGGATCGTC GGTTTATTCA ATGTCCTACA TGACGCTTCA    1200

ACCACATCAG CAAGCATCAC CACCGCAGAC TCCTTTGTGC GAGGAGGAAC CCCCGGATAG    1260

GTAAAATCTT CCCGATCAAC ATTCGGATCG AGTTCCCGCG CCTTCTCGTA AAAGTATGTA    1320

ATAAGACTAT TACCGTGATG CTCTGCAATT ATATCGATAA CCTCCTGAGG TAAGCGGAGT    1380

TGATGTGCCT TTTCTACCCC CAGCTTTACA TGACTCCGAA TTACCGTTGC AGAAAGCCGT    1440

GGATTTAAAT CTAAGTGTTT GCTATCGCCC GTTTGGTTTT CTACAAAGTA CTCACCGTTT    1500

TCCATTTTTC CAATGTCATG ATAATACGCG CCAACTCGCG CAAGGAGCGA ATGAGCCCCA    1560
```

ATGCTACGAC ACGCATTTTC TGCAAGAGTG GCAACCATCA TGGTGTGATT GTACGTACCT   1620

GAAACTGTAA GCAGCATTTT TTTCATGATA GGAACGTTGA GGTCCGAAAG CTCCATAAGC   1680

CGGAACACGG TAGGAGCATT GGTGAGCGCT TCAAGGATGG GCAGAAGGCC TAACACCAAA   1740

ATGCCGTTGA GAAAGCCACT GATCGCCACG CCTGTAAGGA GGAATATTGC GTCAGTGTAC   1800

GCATGCGGAA ACGCAAACAT GAGCGTAGCA GnCAAGGAAA GGCTGAGCAA CGGCAAGGAC   1860

ACAGGAACTT TTAACAATGT CGAGCCGAGA GCTCATAACA CGCATACAAG CAGACGCCGA   1920

CACCCCAGAG AGGAGCGCAA AAAGCGTAGG CTCaGTATGG AACTGTGAAG CGATGAGCAC   1980

TGCGAACGCA ATGAGAAAGG AACTAGTGAC GGCACTACGA TGGGAAACGA GCGCGGTAAC   2040

GAGCATGATA CACAACGCaG TTGGCTGAAA AGGAATGCTA TCcAAGCGGT GCAGGGAGCG   2100

CAGCTATCTT TGAAAGAAAA AGTGTGCACA GGTATCCGGC AACGCTGGTA TAGAGAATGA   2160

GTAACTCTAC ACGCAgTTTA AGAGGAGGAT GGGCCATCCG TTTACTGAAC AAAAAGAAGG   2220

CAAGCAGATA CAAAAAGGCC AGTAACAGGA GACTGCTTAC GAGCAGGGAG CGATCGACAG   2280

ACAGTTTAGA GTGTGCAAGT GCCTGCAATC TTGCGTAGTC AGTGGCGGAT ACGATAAAGC   2340

CGCGACGGAC TATAATTTCG TTTGGATGAA TACTGAGGGT GACCGGTCGT AACCGCGCCA   2400

ATGCGTTGCG GACATGTCGT TCACTTTGAA TAGGGTCAAA GACAATATTT GGACGCAGAA   2460

AGGGTCCGAG GGATGAAAAG AGGAGCGCCG CCTGCGACGT AAGACCGAAA TCGGAAGCCA   2520

GCGCATGGAC GCGCGCGGCG AGTTGATCGG ATCTGATGAG CGTTTCAATT GCCTGAGTCC   2580

TCGAAAAGAC ATCCCCTGCG GCATTTTCCG CCTCATCCGC ACTCGCATCG GGTGACAGGG   2640

GTACAGACGC AGAAGGGAGA TCTCCTGCTG AGGCTTGCGA CGCCACGGGG GCCGACATAT   2700

CCGACGGGGC GTGGAAGGGA CGTTCCTCAC TGATAGTAAT TGTGTGGGGG TTAAAATCCT   2760

TAAGCGCATG GTCGGACAGC TGCACCACAC CTTGCGCGAA GATACGCGCG AGCACTTGAG   2820

TTCCCACACG CAGGAGGGAC TCAAACGTGT CATCGTCAAG CTGAAGCAAG GATCGCAGCG   2880

TCTGGCGCGA AAAGTGAACA AATTTCTGCT GCAGCAGGTG CACGTGTGCA GACGCCCGAT   2940

CGTGCGCAgC GTGGaGTGAC CGCTCCTCCT CGTAGTCCGC TGCTCCTCCG CCTCCGTCCG   3000

ACGAGGTATC CAGAGCCATA CCAACGCGCG CTTTCTGCAA CGCATGACAA AACGCCTGGT   3060

ATGCGCGTAC TTCAGCCTGT TCCAGATCGA GCCGACGCTC AAAAACAGCA GGAATTTCCT   3120

TCTTCCTGCG AGCATACTGC CcGCTGGGTA rCCAGCTCAT CAGTAAGGGA AAGAAAGmCA   3180

GGAGAGACAA CgTTCCgcTC AGtACACGCC CTACCGCAAA TCAGCAAGTT CAGTCTTGCA   3240

GGGTCTTGCT GTTCGCTGAT GCTCACCGCC TTGGCAATGC TGAGGACAAC GAAGGAGAGC   3300

```
GCCAGATTGA GCGCGCGCGC ACCCgGCGGC GAAGTACGTG ACACAGCGTA TGCCACAATG    3360

CGTAAGACGG GkTGGTCCTT TCTTCCTCAT GCGCACTTCT CGCGCCAGCG AGCrTaACAC    3420

GCCAGCGGTA ATCTGTCCAG CAAGGACACA CGGACGCCTT GCGTACCCCA ACCGCGAGCC    3480

TTGACAGAAC ATACCCAAAT ACCGCACCAT CGGCCTCCGC AATGAGAAGG AGTGCGACAG    3540

ACCGTGAAAG GATGCGCCGT CACCATCGAC CAGGTCTCAA AAGCATACGG TCACTGCCTC    3600

GCCGTTGACC GTGCCACCGT TCACATTCGG CAGGGAGAGT TTTTCTCCAT CCTCGGTCCT    3660

TCAGGCTGCG GAAAGACCAC GCTTTTGCGT ATCATTGCAG GGTTTGAACA GCCGGACTCA    3720

GGAGACTTGA CCTTCGACCA CGTGAGTGTG CTCGGTGTTG GTGCAAATAA GCGGAGGTCT    3780

AACACCGTTT TCCAGTCGTA TGCCCTCTTT CCTCACCTTT CCGTGTACGA GAACATCGCC    3840

TTCCCCCTCA GGCTCAAACG CCTCTCAAAG AACCTCATCg CGAGCGCGTG CACGAGTACC    3900

TTCACCTGGT ACAGCTGGAC GAGCACCTGC ACAAGAAACC CCATCAGCTG TCAGGTGGCC    3960

AACAACAGCG CGTCGCCATT GCCCGTGCAC TCGTGTGCGA GCCAGGGGTG CTCCTGCTTG    4020

ACGAGCCGCT TTCTGCCCTG GATGCAAAAC TTCGCTCCAA TTTGCTCATA GAGCTCGATA    4080

CACTCCACGA TCAGACGGGC ATTACyTCGT TTTTATCACC CATGACCAGA GCGAGGCTCT    4140

GTCCGTCTCC GACCGCATCG CCGTCATGAA CAAAGGAAAG ATCCTGCAGA TCGGTACTCC    4200

CTACGAGATT TATGAGCAAC CTGCGACTGA CTTTGTCGCT AAGTTTATTG GGGAAACTAA    4260

TAGCTTCCTG TCAACTGTCG TCTCCTGCAC CnCCATTGAA AACGAAGAGT TTATGCTCAG    4320

TCTCCAGGTT CCGGAACTTG ACCnTACGCT CACC                               4354
```

(2) INFORMATION FOR SEQ ID NO: 63:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21948 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

```
GATACTTCCC AATGGCACTT TcsGGTCGCt GcTTTTtCyT CACgTTaACA GCGAACGTAT    60

TGATTTTAaT ATCCACCTGC CAAAAgGAGG TTCAtTACAG GACTATGCTC ACATCCGCmA    120

CACACTCAGC CGCAGCGTTG CGCACTTCTA CCGTCAGTGC ACTATTGCTC ATACGTACGT    180

GCAGAACTGC CCACGCACtG CCACTCAGGG CAACGCGCCA ACACATTCCT CACCCCCCTG    240

CACCGGCGTA CGAGAAGAAC CCGCCGCTCC cTGCGCGCAC ACACCCCGGT ACGAATCCCT    300
```

GTTCCCTCTA CCCGTGCAGC ATGCGCACCT GCTTCCTCCG TCACCTCCTC ACATCTCGTG   360

CGAACACGCG CGCGATTGCA CTCACCCAGC CCCCGCTGCC GAAGGAGATG CGCCTGTGCA   420

CAACCATACC CATACAGGTG CATTCAAAGT ACTCGGACAG GTAGCAGGAA CATTCATCGC   480

CGTAGAACGC AACAACGCTC TCTACCTTAT CGATCAGCAC GCAGCACATG AACGCATTAT   540

TTTTGATACG CTACAGCGGA ACCTTGGCAC TGCACAAATA CTTCTTATTC CCTACCACAT   600

TCACCCACGC TCGGATGAAG AGGCGCGCAT CATGCACCGC GCCTGCACAG AACTTTCTCC   660

TGCAGGATTT CGATTTCACG AAGAACCAGA CGGTTCGTGG CACGTAACTG CGGTGCCGCT   720

CCACTGGCGG GGGAGCGAAG AGCAACTTGC ACACGATATC CTCTACTCAG GAAAAAACGC   780

GCACGACATC CTGCGCCACG TCCTCGCTAC CTGTGCCTGC CGGTCTGCGT GTAAAGACGG   840

CACCATCCTG GATGACGCAA CGCTCCACTC GTTAGTGGAG CAGGCTTTTG CATTACCACA   900

ATCGAGGTGT CCCCACGGAC GGCCCATTTG GATTGTCATT GGCCGAGACG AATTGTTCAA   960

ACGGATCAAG CGCACGTAAC GCGCTGCAGA TACGCAAAAA GAAGCCTGCT ACGTCTGCGC   1020

TCTCCGCGTC GGCACGGGGA GGTGCGCCGT GTGCACACAA ACACCACCAG TGAGAGGATG   1080

TACGGCAGCG CAAACAGTAC ACCGGTGGGC ACCACGTGAG TGCCCTGCAA TAmGTcaCAC   1140

ATGTGTTCAA TACCGGAGAA AAAAATCGCC GCCGGCACAC ACCACATCAT CCGCTTGCGT   1200

GCAAGAAAAA CAATTGCAAG TGCCGTCCAT CCTCTGCCTG CAGCCATCTG CGGGGTGTAg   1260

TACCGACACG CAATACTAAC AGTCCCCCCG CACACACCGC ACACGCCT GTaCCGCCCA   1320

CGACACCATC CGATtACGCG CCGCGTCAGT TCCCCGCACC TGCAAGGTAA CGCACCTTCC   1380

CCrkAGTGCA TAAAATTGAT ACCCACGTTT GTAGAGTACA GATACAGGTG AAAAACCCAC   1440

ACCAGTGCAA AGGCCACCGC AGTCCCCCAC AGGGGGTGAG GTAAACGCG GGTATGTGCA   1500

AGAGAAACAT GAGTGAAAGA GACACCATGC GCTGCCGTGT CCATCTGCAT CGCAGAAGCT   1560

GCAGCGCGTG CAAACATGCT GGACGCACCA AATGCGCTCA TCCCCATTGC AGAAAAGTGC   1620

ACTGCTATGC CCGTTAAAAA CGGATTTGCC CGCATACGCT CCGTACCCAC GGCCACAAAA   1680

AATAAACACA GCGGCACCAC ACACACGGTA ATACCCAGTC CACCCCAATA ACTTCCCCAT   1740

ACCAGTGCGA AAAACGCTAT GCAAAAGGAC GAGAAGGTAA TCACCCCTTC CATAAAAATT   1800

CCCAACACTC CCGCGTATTC TGTTGCGAGC GCTCCTGCTG CAGCGCATGC AAGCGGTGct   1860

GCGCGATGTA ATATTGCTAT CACTGTGGTG CCTATCACTC CCATCGAGAC CGCCTATGAT   1920

GTGTATCATG TACAGAAAGC GCATGACGTC TGCGAGTTCT GTGCTTTTCC CCACGAAAAC   1980

AAAAAACGGT GACAAGGAAT CGATATACCC GGCGTGCGCC TCGCCGCACT GCATTCCACG   2040

```
GTGCGGACCA TTGTGCAGAA ATAAGCAAAA AGATCGCCGC CTGTAAAAAT AGCACCACAT    2100

TTACCGTCAG gTGCGCACCA AGTACTGCAG CTTCAGAGGC TGTCTCCATC CACGCGAAAA    2160

AGAACGCAAG CGGTACGAGT ACCGTAATGT GTGCATGGGC AATTAaCGCG TGCGCTAAGG    2220

CTGCGTAACC CATCCCCACA GAAAAACCCA CATAGCAGGT GCCAAACAGC CCAACTACAG    2280

AAAAAAATCC GGTAAGCCCA AACAGCGCCC CTGAAAGCAC CATTCCCCAC ACATAGGTGG    2340

CCCATACGGG AAACCCTACA AAACGCCCAA ATTCGGGGGC CTTTCCGCAT ATGCGAAACT    2400

GATATCCTAC GCGGGTGTAC GAAAAAAAAC ACCCAACTGC GAGTGCTACT AAGGACGCAT    2460

AGGTCAATAC GGCCGGCACA CCGAACAAAG ACGTCTGTTG CTGCAATATA AAATGCGAAT    2520

GAACCGGCGC AGTTGCCAGC AAGTTCCCCG CAGaTCACGC GTAACCGTTA TAATCAACGC    2580

ATCGATGAGA GGCACGCATG CGGTGGATAA CAAAAAGGAA GTAATCATTT CGCTAGTTGC    2640

CAGCCATGCT TTTAGTATCC CAGAAACACA GGCTAATATC CCCGCGACCG AcAGCGCACA    2700

GAGGAGCGCA ACACTCCATT GCAACAAAAA GCCCACACCC CAGTACTCAC GGAGCAACAA    2760

TGCGGTGACA AAACCTGCAG CATAGATCTG GCCATCACCA CCTAAATTGA TCATTCCTGT    2820

TTTTAGCGCG CAnTCGCCCC CAGTGCCATA CAGACAAACA GTCCTGCTTT GTGAAACAGG    2880

GCACGTATGT AGCCACGGGT AGAAAAGGT TTGAGAAAAA ACGCTGCCAA AGATACGGAT    2940

GGATTTTCCG AGCACAGAAC AATCACAGCA CTCATAACTG CAACACCGAG CAACACTGCG    3000

ATACACGAAT TGATCACCCG TTTCACGTAT GAGAATCCTG AGACGGAGAC GGAGTGCCTG    3060

ACACTTCAGC ACACAACGTA CCTGCACGTA GCAAGAAACG TTCTGTGCAC AACGCACGCC    3120

ACTGTGCCTG ATGCTGTTCT CGCGCAAGGA GCACAAGAgc AGTTCCTGCC TGTGCTACCT    3180

GGCGCAGACG TGCAAGCAAG CGCTGTTCAC TGGCGCTATC CAATCCTTCT GCAGGTTCTG    3240

CCAAAATGAG AAGACGTGGA CGCGTTGCAA GctCACGCGC TAAAATAACG CGCTGCAACT    3300

GTCCGCCTGA AAGCGTACAG GCAGGCTGCA ACGGATCGCA GTAAATTTCT TCTTCTGCAA    3360

GAAGACGAGC AACAAAGCGC ATCTGGcgCg GCACACGCGT GCGCCACGTA CGCAACGTGT    3420

AGGGAACGAG CAAATCAAAA AGAGTTAACT GCATTGAGGC ACCGCGCTGT ATGCAATTAG    3480

ACGGCACACA CGCAACCCCG TGTGCCCGCA GCAGCGAGGG CGTATTGCGC TGGAGGGGGA    3540

GACACCACAC CTGATCGTGC TCCTGCAAAA GAATATTCCC GGTGCAGTGC GTACGCGACG    3600

CCCCAGCGTG CATATCACAC AGTATATCTT CCAATACGTG CACACCATCT TCTGGCGTAC    3660

CGACTATCCC TATGATAGCA GATGCGGCCA CAGAAAACGA AATATCTGTG AGCGGAACGT    3720

CTGCGTGTTT ACTCACCTGC AGCGACTCAA CGCGCAACAC CCAAGGACGA GCAGAAGATG    3780
```

```
TGCGCGGCAC AGTTGCGCAC GACTGGGTAT CTGACAGAGA GGAAAAAGAA CTTACGGCAG    3840

AAGAAGTCAC CGTTGATGCG GACATGAGCG CACAGGACAC TTTCTGAATA CATTCATTCA    3900

CCTGATGCGC AGAACAGTAT TCGTCTAAAA GATCCGTACG CAGAAAACTG CACGCTTTTC    3960

CCCCTTCTAT CAAAGAAATA CGCTGTGCCC ATCGCAATGC ATCAGCAAAT CGGTGCGTCA    4020

CTACTATCAC TCCGCCACCA CACCGGGGCG CGTGCGAAGA ACGCACAAAA AACTCTTCAA    4080

GATGAGAAAA GAAAACCGCA CGCGATTGCG CCGGAGCACA CCGCGGCTCA TCCAGGATGA    4140

TGAAACGCGG ATTGCGAAAC AATACACAGA GCAACGATAC AAAAAACCGC TTGTCTGCAC    4200

TCAAACATGC AACGTATTCT TCCTTCTTCA AGGGCATACG CCACTGGGCG ATAATGCGAT    4260

CTATGCGTTC TCTCACcTGT GCACGGCGCA CCCACCGCAC GCCGGTgAgT GCAGCACTAC    4320

CCATCACTAC ATTTTCAAAT ACTGTTGCGC GTTCTGCAAA TACCGGTTGC TGGTGCACTA    4380

TGCCAATTCC TGCACGGAGC GCATCGAAGG GTACGGAGAA GCGCTGCTCC TTTCCATCCA    4440

GACGGAGCTG CCCATGCGTC GGCACGCAAA AGCCCGAAAG AATATGCGCA AcGTGGATTT    4500

TCCTGCACCA TTTTTTCCCA ACAACGCGTG AATTTCACCG GTAAAAAAGG AAAGATTCAC    4560

ATnGCTGAGC ACGCTGTGCT CAGGrCsGTC CGTCTCACGC GCGCCCGAAC ACGGGCCATG    4620

CGCTGTGTGC GCATCGTCGA CTGCGCGCCT GCCAGGGTGA CCGAACATAC CCCAGACCCC    4680

GCGCTTTGAA CGCGGCATCA CGCGCGGGTA GGTTTTCCCA ATATGGTGAA GCGAGAGCAC    4740

ACCGCGCGCA GATGCCCTAA CGCCGCGCTC AGCTATCATC AACGCACCGG CAACGTAAGC    4800

TCACCGCTTT GAATACGCCT GAGCAACGCA GACTGCCGCA CACGAATCGG TTCGGGTACc    4860

GTTTGCAGGT ACAAGGGATC CTCTTCAATG AAACGTACGT ACCCGTCTTT CACCCCCAAT    4920

GTCCAGGCTC CTGCAGATGG CAGTTCACCG CGAATGCAGC GCAGtcTGCT cATACGCAAG    4980

ACGCTCCTGT TCCATAACGG AACTGCCAAC TACGTAGCCC GGTGCCCTCG CATAGCCGTT    5040

ATCGTCAAAC CACGAAACAT AAAAACCGAG CTCCCGCGCG GCCGCAAGTA CTCCCTGATT    5100

CGCACCGCCG CAAATTGGCA TCATAACATC CACCCCTTCG TGAAAGAGAA TCCGTGCGAG    5160

GTCTGCACTT TTTGCAGCGT CATACCAGTT CCCCACCACG CGCACATCGA CTTCAAAGGC    5220

AGGATCTACT GCACGGGCAC CTGCGAGAAA GGCAGGAATA ATAGTCTGGG TCATCACCGG    5280

ATACGACTGC CCCGCAATAA GACCGATTTT TTTATCTGCA TTTGCAAAGC GCATAGCACT    5340

CGCACTCACT AACGCGGAAA GGTGTCCTGC AAGGTAGGCT TGCTCCCACT GGTTATAGCG    5400

AAAGGTAATC AGCGAgTGCT CtGCGGCGCG TAGGCATCTA GAACCAAAAA CCGCTGCAGG    5460

GGAAATTGAC GCAAAATAGG CTCAAGGACG TGCGGGAGTG CAGGGTTGGA AGACACAATC    5520
```

```
AAACGATAGC GCTGTTCTGC AGCAAGATGC GCCAACTTTT CGCGCCAGAG CGCCTGGTTC    5580

GGCCCCGCTT CGATGATATC AAGCCCaATG TGCGCCCTGT CGCGCGTTCc TGCGTAACTG    5640

CACGCTCAAC ACCGTCACAC AACATTGCAT ACACAGGACT GTCGTGACGA AAACCTGGGA    5700

CAAAAACGGC AATAgCACCG CGCGCTCATC TTGCACCGCA GGCCTACACG AAAAGCAAGT    5760

AAACACTGCA ATGAGCGCAC TGAGAACACA CACCGCACCG TTCATAACAC CTCCCCCCAA    5820

AAATCCCTCT CTCGCGTAGG GTGCACCcTA CCGGCACCCA CAGCCTGAAA AGACCAGAGC    5880

ACTACTCCTC ACCCCTGCCC CCAAACGCAT TGCACCACCC AGAGAGAAGG AGAAAGACTA    5940

GTTTTTCACG CGTTCTTGAA AAACGTAGCA TCCGCTAAAC TCTGCGATCT CACGGATGCT    6000

CTTCATCCAA TATACCGCTT CTGTACGATT AGTAAAAGGA CCGACACGCA CACGGTGACG    6060

CAAGCCTGCA CCCGTGCGTT TCGTGAAGAT TTCTGCCTTC ATGTGTCTAG CTGCAAgCAC    6120

ACCCCGAGCA CGCTCGGCGT TGAGCTTACT TGAGAGCGAA GcGGCTTGCA CCCAGAAAAG    6180

GACAGAAGGA GCAGCAGGCT GCGTGGACGC ACCGCGCACG CGCGCATcCC GCTGTGCAAC    6240

AAGCGGAGCA CGGGTACGAG TAGCGGGAGG CGACTTGGCA GACGCTCGGT CACTCCGTGC    6300

ATGCTTTGAC GCACCCTTTG ACTCTGCAGA TTCCCGAGTG TCAGAGGCAG GAGAAGTTTT    6360

CCTGTCCTGC GCCGCGTCTG TGCGCTCAGC AcGCGCCGGA GGAGAAACAT CAAGACTTTT    6420

CGCACGAGCG GTAGGGACAT CCTTTACCAC CGTGAGATCA GGAATTGCCC TCTGTGTTGG    6480

CGCAGGAGTC TTCCCCAGCT CAGGAATTTT TTCAGGATTT TTTAGCCATA AGCTCGGATC    6540

CACTGACGGG TGCTCAGGTA CGTCACCACG CTCAACATAA GAGGACACGT CGGTCACTCC    6600

TGCAGAATTA GAAGAATAGG TAGGAGAGTA CAACAAAAGC GCGACGCCAA AAATAATGAG    6660

CATAAACACA CTCAGGGAGA TGACAATCCA CAAAATTCTC TTCTGTTCCA TATTTACTCG    6720

CCTAAAACAC CCCTGCGCGT CAGAATGCAA TCGACTTTTC GCGCCAATGA CGCAACACGC    6780

CCACAGTTGG CCACCGTGTA GGTATCGGCG TTTTGCGCGA TGCTTTGAGC ATAGAAACCC    6840

TTCTGCGCAG AAAACCTGCG TAAAAGGTGA GTAAAAGATA CTCGTTCCCG TTTCTTACAT    6900

CGCCACATGC GCAGTATGCT TGGCGCCCAT ATGTACAAAA CAAAACTACA GGCTTGTAAC    6960

AGTTCTGTTT TATGCAATGT CGGCGCATTC AGTACTATCG CTTTTGGACG CGCAsCTGCG    7020

CGCGCGCTAT GTCTTCACAC AAAAGACGGG TAACTTTCGG TAACAAAAAC TCTTCGTGCT    7080

GCTGCAGTAG TTTAGGCTCA GAGAACAGTA ACACACCCAA ATGCGCAGAG TGCAGGCCGC    7140

CATCTTTACG CCGTAACGAA AGGCCGCGCG CAGCgCAGCC ACTTGAAACC GCTCTACTAT    7200

AGAGTCGCCA TAAGATTCCA AAAGTTCAcG CGTGCGCGCA TCCGCGTCAA TACAGTAACA    7260
```

TCCCCGTTCT GCAAGCAGGC GCGAGACCAC ATTCTTCCCC GCACCACTTC GACCGATGAC   7320

ACCAATTAGT GGACAAAACT CGCGCACAGC GAGAGCGTAA CGTCAAACAC GCACCGTCCT   7380

CAAGTGTTCA GAATGTGCTG AACACCCGAT ATTCCTGCGT ATCTTGCGCT GACTACCCCT   7440

GTGCCTCACG CGCAGCACGC AAAAGACGCT TCTGTGCATA CTGTTCGTGT CTGGCACGGT   7500

TGGGAATAAA CTCACGCTTG AGACGACACA GCCGTTGCGC TGCTGTTTGC ATATCCGCAT   7560

CAAAATCCAA GGCAACACAC GCGAGCGCCG CGTTTCCTGT CAGTTCTGCG TCATGAATTT   7620

CTGGTAACAC AAAACAACGC CCACTGACAT CAGCCTTAAG CTGCAACCAA CGACTATCCT   7680

TTGCCTGTCC ACCTGAAACG GTGTATACCG GTTGCAACTG GGTCACAGAC TCCAGATACT   7740

CCAGACCTGC ACACACCTCA AAGGCAAGAT CCTCTACCAG CTGTCTCCCT TCTCCCACCA   7800

CGGTGGGGGG ATACGCATCA TGCAGCTGAA AAGGTAACGC CATAATTTGA CCCATCCGTT   7860

CCCCAAACCC TTGAGCATGG GTGCGCCGCA TATAAGAGGC AAAACGACTC CCTGAATCCG   7920

CAATTAAAAA CGAGACATTC CAAAAATCTG CTCTCAAAGA CGGTAACACG CGGACCCTTG   7980

CATCCGCAGA CCTGATCGCA GGAGGCAGAC GAACACATAC ATTTAACCCC TCGCTGGATC   8040

CTGCCCGATC GCATCCGCTC CCTGCATGTA ACGTATTTGT TCCAATGAGT GCTGCCGCAA   8100

AGTCCGGCGC GCCGCAGACT ACCGGTATTC CCCGATAGGA GGCAACAATT GAGCCAGGAG   8160

CCACGAACGG CGCAAAGAGC GTTTCAGGTA ACGCACACGC GCGTAAACTC TCAGACGTCC   8220

AGTACGTCGG CATGTAACGT CGCTCTGGTA ATACCGTAAC CGCACACCCC GTGAGTCGAT   8280

ATATCAAGTA TTCGTGAGAA GAAAGAAAAA ACTGTACATC ACGCGCACAA AAATGCAATC   8340

GCTGCAAAAG GAGCAAAACC TTTGGCAAAA AAAGAGAAAC ACCGCACCGG GGATCAGACG   8400

CCCCCGCTTG ATTCCACAAA ATCAGTTGAT CCTCGGCATG ACTCTTCTTG TGCACGGCAA   8460

CGACGCTCGG TCCATTGCCC GAAATAGTAA TGGCAATAAC GTGATGCACG GCACGCAACC   8520

GCTCAAACAC CGTAAAAAAT GAACGCACCC AATCCTGGGC CTTCACCGGC TGAGGAAAGA   8580

ACACGCGCTG GTACTGTAAC ACCTTTCCAT CTTGGGAAAT AATCGCCGCT TTTAGGGAAG   8640

ACGTACCGAT GTCCGCGACA AAGATCCCCG AACCCATGTC CGTTGCGACG CTACCCTATG   8700

CGCGCGGTGT CCCCGCAACG GACAAGCGAG CGTGTAGGAG TCGCTCGATG ATACTGCGAT   8760

CATCCAAGAG ACTGCTCAAC GATTGATTAT GATGAanTGC GCGcaTGACG CGcGCGAACA   8820

ACCCTGACAC ATCGGTTTCA GTATACCACT GCTTGTGCGA AAGCTGTGTG TGGAACACCG   8880

CATTTGTGCC AATAATCCGA GAGAAATACC GTTTCTCGTA CGCTTCATCA AATAACTCAA   8940

GTGCATTTCC CGTAAAGAAA GGTAAACTCA CTGCCGCGAT CACCTGCTTT GCTCCGCGAC   9000

```
TTTTCAAAAA TTCCATCGCC TTTAGCATCG TACCTCCGCT GCCAAGCATA TCGTCAGCAA    9060

TAAACGCCGT CTTCCCCTCC ACATCGCCGA GCAAGTTAAT TTCTACAATA TTGCTCTGCT    9120

TTGCATTTTG CGCGACCACC GAATAATCAC GCACCTTATA AATCATCGCG AGTGGCTTTT    9180

TTAAACCAGA AGAATAAAAT TTATTCCGTT CAACCGCCCC GCTGTCCGGC GCCACTACTA    9240

CAAAAGGGAT ATCGGGGTCA GAAAGATTTT CAATCTTTGC CAACTCCCGG ATAATCTGAT    9300

AACTGGCGTG TAAGTTTTCA AGCCGCGTGC GATGAAAGGC ATTTTCAATC TCACGTGAAT    9360

GCAAATCAAG AGTGACAATG TGACTCACGC CAAGATACTC ATATACACTC CCGAGCAAAC    9420

CCGCCGTCAG TCCCTCACGT CCACACTTTT TGTGCTGACG GCTATACGGA TAAGTGGGTA    9480

AAACCAAGGT AACGCGCCCA GCTCCCGCGT GCCGAACTGC ATCTATGGTC ACAATGAGCA    9540

TCATCACGTG ATCATTCACG GAGAATATTT TTTTACTTTT TCCGTTATTC ACCAGGACTG    9600

GTTGATGATT TTCTACATCT TGGAAAATAA AAACGTCCTT GCCACGAATA CATTCATTAA    9660

TTTGCGTTTT TAACTCACCA TTTAGAAAAC AGATAAACTG TGCATCCACC TTAAAATGCG    9720

GTGGATTAAA ACGACGTACG TCATCGTGCG CACACAACTC TGTGGAAAAA AGGTCCCGGT    9780

AAAAGTTCGC ATCTCGTATC ACCGACCCCG AGTCAAGACC GTAGCGGTCC GTCAAACGGT    9840

CCATTCTCTG GTGAAACTTG CGTTCACACA CACGCGTCAA ATGTTTGATA GTTTCGTCCG    9900

CGAAGTGCTC GCCACCAGGA CAGGCGACGA TCGCCAAATC AGTAAACCCT GAACATCTCA    9960

TGCAATCTCT CCACACTTGT CAACGCGGAC TGGACAAGCA CCGTCTCAAC CGCACGATAG    10020

GGAGCCCATC CGCGCAGGCT AATACAACGA AACACACCCA AGCGAACACA CCTTGCATAC    10080

GCAAGGCACA GCGAACGCAC ACGCGTCATG CGCACAGGGC AACACTTTAC TTACTTATGA    10140

TAGTGATTTC TACCCGTCGG TTTTTTCTAC GACCATCCTC TGAATCATTT GGCGCAATAG    10200

AcTGctGCGC ACCACAACCG CGCGTATATA CATGCGCTGC ATCCACAACA CCTAATTCCT    10260

GCAGGTAACG TGCAACCACA TCAGCACGCT CTTCAGAAAT CCTCTGTTGA TCCTGCACAG    10320

ACCCCCGTCG TGCCGCATGT CCAGACACCA ACAACTCTCG ATCGGGAAAC GCGCGCAAAA    10380

GTTCTGCTAT TTTGCGCAGC TTCTCGTACT CAGAAGGTGC AAGGGATGCA GAGTCTGCGT    10440

CAAATTGAAC ATTTTCTATA CTGATGGTTA CTCCCTCTTC TGTTTCACGC ACcTTCGCAT    10500

CAGGCATATG CAAGTCCTTG AGCGTTTCTT GCAGTTCCAC CACCGTGCGT GCAGGATCAA    10560

AGCGCTCAGG TGCAAAATTC TTTGCCGTCG CAGTACCCTG ATATCTCAAA ACCGTTCCTC    10620

CACTCAGGTA TAAGAGAATG CGAAACTCAT CGTCGTACTC AGCTATGTTC CCAAGTTCGT    10680

TATCCCAATA CAAATTTTGC TTAGAAACGC CCGTAGTACG CACCGGATAC ATTCCTTCTT    10740
```

```
TTGCATTGCG CTGCACGCCA TGTGTCCGTT TGGGCGACTC ATAACTCATA GAATACGCCG    10800

CAGTAATGTG ATGATAGCGC CGACTACCAC GCTGAACTTC ACCACGGTAG GTATACGACA    10860

CGGTAAACGG CACAATGAAG GGCGTTTGGA TACCAAAGCC GTCACGCAGA TCATGCGCTT    10920

CTTCTGCTTC ATGCTCCCAG GTATCTCCAA CTTCGATATC ATAATCAGGA AACACCGGCA    10980

CATTCCGTAC TACCGGCATG AAGAATGAAC GATCAATATC ATACACACCA AATGCGTCGC    11040

GCCAGAAAAT ACTTTCGTAG TGCCTCCCCC AACGAAACGT ATTATTGGGA CTTTTCTCGG    11100

ATGTCATGAA GTGACACACA TACCGCGCCG CATCAGGCGC CGACCCGTGT GCAACAckTA    11160

CTTCGGAAAC ATGCACCGTG ATTCGATTCG TAATCTCCGC CGTGTGAGCA AGCGTATCGT    11220

TCACAAACAC ATCCTCGCGT ATCAGCGAGT TGATACGGTG CGTATCCCCC TTACGAAACT    11280

TGTaGCGCAA GcgCAGAGGG TACGCCGCAC TTGCCGCGCA CCAGACTGCA AGAACACACA    11340

GCATTCCCTT CCTAGAAAAC ACACTTCCCA TCATACACAC CGAACGGACC TATCCGTCAT    11400

ATTCCGCACG AGAAGTTGCA CCCCACCCTT TCCTCCTTTG AGAAGAGCTG TGATAAAAGG    11460

ACGCCCTATC AGCACCTGct GCGCACCCAA CTcCCstGct GCGCACAAAT GCGCATAACT    11520

GCGAATACCA CCGTCTACCC ACACTTCACC GGCACAGCGC GCTAACTCTC CACCGTATTC    11580

AAAAAGAAAG TCCgCTGTGC TTCCGCGCGC AGTTTCAACC cTGCCTCCAT GATTTGATAC    11640

GATAGCGACA TCCGGTTTCA ACTCTCTGAC TAATTCGACA TCGCGCGGCG CAAAAATTCC    11700

TTTCACCACG ATAGGAAGTT TTGCAAAACG TCTGACTGCA CGAAGGTGsG TAGGAGTCTT    11760

TTTCTCTAAC TGCACTTTAT CTCTCATGGT CACAATATGG TACGCATCGA TATCCACACC    11820

CACAAACTCC GCAACGTCCC GACCCCACTC GATACGTTCA AAAATTTTTT TGTTCACATA    11880

CGGTTTGATA AACACTGCAG CCTTCTTTTT GAAGGAACGC AACGCAGCAA TACCCGACTG    11940

CAGTTTGATG TCCGGACAAC CATCTCCCAC ACTCAACAGG ACACCcGTTC CGGACACCGC    12000

TTCGATGAGC CGATAATAAA ACGAGACCTC GTCAGGATAT CCGACGTTTT CAACTGCACC    12060

CGTCATAGGC GCAAGACGCA CGACCGGTAG TTCATGCCTT TGCGGAACAT ATTTTTTCCA    12120

TGCAAGGCAG TTCGCGATAA AATTCGCACT GTTAAAAACA CCCCCCATCC CTGGTAACTG    12180

ACCACGACAC CCGTATCCAT CACAACGGGC ACACAGGCGA CACTTGTACT CTGAACGCGC    12240

AGAACGCACC ATGCGATCTT CCTTTTCCCT AAAAAAGCAC TGCCCTCTTT ACTCTACCAC    .12300

CCCATAGCGC TTCAAAGCCT GTGCAAAACC TGCATCATCA TTAGATTCGG CAATATACGT    12360

TGCGTGTTTT TTTGCTTCCT CATGACCATT ACGCATACAA AAGGAAACGC CCGCTGCTTT    12420

GAACATGACG ATATCATTCC TCTGATCACC AAACACGCAT ACTTGTTCCA AAGAAATGCC    12480
```

```
GTAGTACTGG CACAATATAT GCAGGGCATT CCCTTTATCC ACTCCCTCGG GCATTACATC   12540

AATCAGGTGC GGCATGGAGT GTTCACAGTG CACTCCATGC AAACTGCCAA TAAACTCAGC   12600

TGCTGTTTCA AGATCTGAGG GAGAAAAATT TTGCACAAGG ATTTTCATAA CCTGTCGCTG   12660

CCGATCGCCT GATAGTGCTT CTATGGGAAA GATAGGGATA AGCGCGCTCC CCCCACGACG   12720

GGCCAAACGA TTGTATTCAT GGAACGCCCG TATACGGACA CTGTAGTCCG GTGCATATAC   12780

ACTGTCAGAC GTATAAACTA GGTAATCTAG GCGATGCGCC ATGCCGTATT CCAAAACACG   12840

AGCAAATGCT TCAGCAGCAA AACATTTTTG AAAAATAGTT TTGCCCGTCT TAATTTCACG   12900

GATCATTCCC CCGTTATACC CAACAACCGG TCCAACCAAC TGCAACTGCC GCACATACTC   12960

CTGCACCATA GGCAGCATCC TACCAGTTGC AATAACAAAA GGGATATTCT TCTCGTGCAG   13020

ACACTTTACC GCCTGCTTAT TCTCAGGTGG AATGTTATTT TCGCTATCCA AAAAAGTACC   13080

GTCCATATCG GTAACCACCA ACCTAACCTG TGCGCCTGAC ATCAACATCA CCTCTGAACT   13140

TGAACAAGTG CACACCCAAA AAACAACACG GCAACTTCAC TTACTCGCAG ATCCATGCGC   13200

GTCTCCTTGC TTTATCGCGT ACACCTCCAT AGTGTCTCCT AATTGGAACC ATTTGCTCAA   13260

TGCCATCAGC ATCTTCCATT TCAACCCACC AGATTTTAAA TGCTTGTGAG GATAAAATCG   13320

CTCGGGATGA TGACCAGTCA CCACTATTTT CTTTACCGTA AAACCGTAGC GCAAGAGCTG   13380

CTTTTTTACC GTGCGCGCAT CCCATATTGT AAAATGATCG CAAGGACTCT GCGCAAAAAA   13440

AAGATGCGGA AAACGCATAC CAGTGACACC TGCAAAATTA GGGGTAGAAA AGGCAAGAAT   13500

ACCTCCAGGC ACTAAAAGAT CTGCCACCTT CCTGAGTACC GCCTCCAGGT CCTGAAAATG   13560

CTCTATCACA AACCACAGGG TAACGGCGGA GAACGTACAC TCTCGAATGT AAACAGATAC   13620

TTGTTGGGGA CGATTCGTGG CAAAGCTGTG CCGAATCACA AAGTCAAAAC ACTCAGGAAG   13680

TAATGGAAAC GTTGCAACAC AAGCAGGAAT GCAGAGTGTG TCTCGCACAT GACGCACCGC   13740

AAATTCACAG ACATCAACCC CAACAGCATT CCACCCCGCA GCCTTTGCCG CAGACAAGAA   13800

AGCGCCATAT GCACAGCCAA CATCTAAAAC CTTTTTATCA ACTGCAAACG AATTACCTTC   13860

ATCTGCACAA AAAACTTCTG CATATAAGCG TTCGATCTCC TCCATCCTGC GCGCACCATG   13920

CATTCGAATT TGGTCAAAGT CCTCAAGGTA AGTCTTGCCA TACTGTGCCT TGTATTCTTC   13980

AAAAAAATAT GACTCGGAGT ATCGAACTGG ATCTGAAATT ACAAAGGAGA GAAAAATCAT   14040

ATCGCATTCA TTACAGCGCT GAAAGGTTTT ATGTAATGCA CGACCGACAA CATCCACCGC   14100

TACCATTTCT CCACAAAAAG GACAGCAGTG TTTTTTCCCC CGCGCAACGc gCATGaTTTC   14160

ATCTGCCAGA TCACGACACT GCGAATGTCG CGTCACAACA TCGGGTATAA TAATTCCCTG   14220
```

```
CTGCATCTGA GACCACAACT CATGCGCAgc AGGnATCAGC AGATCGAATA ACAGAAAAAC    14280

CCACTGcACT CGAGAGTAAA AAATGATAGG GCGTTGGTGA AACAAGCAAC ACCGCCGCCC    14340

CTGCAGCTGC AGCCTCAAAT GCGGTAAAAC CAAAATGGGT AACTACCACG TCCCAGCGGT    14400

GGAGATGCTC TTTCAAATGA GGCAAAGAAG GGTATATGTG CACCTTACCG TCCGTCCGTT    14460

CAGACGCTTC ACCAGGAACT ACCACTGAGG TGTCAAAACC TAATGCAGCT ATCCGCTCAG    14520

CACACGAACG CGCGCGGCCG TGTGTATCTT CTGCCCCATA CACCACCAAT ACGGTGGTGA    14580

CGCCAGGTAT AGGAAAAAAT CTGCCGTTTG CCACAGGCAG ATCCTGCTCC TTTCTCGTCT    14640

CTGGAAGTGG GATAAACGCA CTATCACGTA CATTCGTAAG AGCGGCCAGC GACCTTCGAC    14700

CCGGACTTTG CAAGACGGGA AATACATCTA TCAAGTAATC CGCATTCAGA CGTCCAGATC    14760

CCCCCTCATC CAGAGCAAGC ACTGGTGCAG TACGCTGAAG TAATTCAATC TCGCACGTAG    14820

AAGTACGGAA GTTGTCGACC ACTATCAGTG ATGCATCCGC AGAACCGTGT CCACTATCAG    14880

TTCCTCAGGA AAGGGAGTAG ACAGTTTGCA CAGCAAACTA CGATCGGGCA CATACAAGCA    14940

ACAGCGCACA CGTCCTTGCA GTCGCAAGAC TAAATACGCC GCCCGATATA AATGCCCCGC    15000

TCCCTGTCCT ATTTTCACTG AGGGTACAAA TACCACCACC TGCCGCGCGT ACGCATACGC    15060

ATCTAAAATA ACATCGGAAG AGAGCGGAAA CGGACACTGC AACGTGCACA CGTACTCCAT    15120

CATATGCTGC GCTCGCTGAA AGTCTTCCTG CGTGTCCACC GTTACTCGCA CATCAGGGTG    15180

ATACCATGCG GCAGCGGcAG GTTCACGCAC ACACACGAAA ATACCCGGGC GGCGATGTAA    15240

GGCAGGTCCT ACATGCTCAC GGTCGTACGC CTCCAAAGGA AGACGATCTG CTAAAAGCAA    15300

CGAGCGCGCC TTTAATATTT CCACTCCGCT GCCGTAGGGA AGACCAGTGA AGGTAAAATA    15360

ATCTGGCTCG TCCAGTTCCG CATAACGCAG GAGCGCTGCA GCAGCTGCTT CGTGAAACAA    15420

AAAAGGGTTA TCTCCGGTAA CCCGCACGAC GGTTCGAATT GGGAACGAAT GCTCAAAAGC    15480

CTTAACTGCG ATACAAAGC GGTGGAGCAC ATCTTCTGCC GATCCTGAAA TACAGTAGAA     15540

CCCATGCGCG CGCGCAACGG GTTCAAAATC TTTTTTAGAA TGTTCATCAC ACGCAAGAAT    15600

ATACGTTTCT GCAGGGATGA CCCGCGTTGC CTGCAACACG TAATGTATCA GCGGCTCCCC    15660

CATAAGGGGC AACAGCGCCT TTCCTGGTAA CCGCGTAGAA TCAACGCGTG CTTGCACAAT    15720

AACTGCAACA CCGGAACGAT CCTGCTCCAT TAAACgTGGT TCCCACTCAC TTACTAAACG    15780

CGCGGCAGTT TTTTTAAAAC GGTATCTATT CACCGCCACC CATTTCTGAG CGTCCCGAAA    15840

GAGACGGTAC GCCTCTATTA GATGCAGACC AGAATTTCTA AAAAATGAAA AAAACTGCGT    15900

GCCCGAAATA AAGGCCCTAT CCTTCTCTAG GAGTGGCGCA AGGTTTTTCA GATAAAAGAT    15960
```

GCGGTACGAA CCATCTGCAC TTGCATCCTC TTGCGGGGGA AGAGCCTCAT AAAACAGTCT   16020

TATTTGCGTA TTAACCTCTA TGCACTCCCC CCACAAATAA GAACGCAATC CAAAATCAAG   16080

ATTTTGCCAA TATGCATTCG CGATGGTAAA ATCAAAACCT CCGCATTGGA TAAACCGGTC   16140

CCGATGGTAT ATTCCAACAA AATCATACGG GTATATCGTA GGCGTATGGT TAGTCGTACA   16200

TTCCGTAGGC TGTGTAAAGA AGTCACTTTT TCGCAAGGCA GGGACAATTT GCGTTGGGAG   16260

TACAGTACTG TGGGAAGAAC ACAATTGCGG CGCAATACAC ATGTGCGTAT TCGTGCGCAA   16320

GATATCCTGT ATGTGCTGGC TCATACCTGC GGATCTATGC GCATGTCGCT CCACAGAACT   16380

AAAACATACA TGACATCAAG TTCTGCAACA CCTAAATTGA TCATTTCTCC AACCGAAATA   16440

CACTCAAGCg GTGTGATAAA CTTAACGAAG GGGAAACgCT CAGCAAGACC GGTAACATCT   16500

GGAGCTGCTG CACTCCGCTC CACCGATACA ATGGGCGCGA TGCCAAGGGT AGTCAGTTGC   16560

GCAAAAAACT CTGCCCGATT CCACCGCACT CCCCGATTGA GTACTACTGC ACCTAAAAAA   16620

GCAGACGCCA CCGGACAGAG TTGACCACCC ACCACCGTGT GTGCAAmATT TTTCTCGTTA   16680

AAAATTATAG GTATAGTACT CATCACACTG CCCACATAAT CCTTCATGTA CTGCACGAAC   16740

GTGCTGAAGA TACACATCCT GTCCCCGCCG CCAAATCTGT GCAAGATCCT GCTGGAAAGC   16800

ATTGCCCAGC GCGTGGCGAC AGTGCACATC TTCTTTGCAC AGCGGAACTC GACCATCAGT   16860

GAAAATAATC ATATCTCGTT TTAGATGCCA ACACGGATAC CTCTCCAAAG GAGATAGATC   16920

TGCAACCCGC CGATCGGGAA GCAATCCGCA CACATGATCA TACTTTTGAA CAATCACCTG   16980

CCCCACACGC TCCTTCCACG TGCGGTAAAA GGGCTCGAGC TCTTTTTCGT TTTCATTCAT   17040

ACGGAAAATC TGTGGCCACA GCACGCCCGG ACATTGCGCA TGTACCTGCA TTGCAAATTC   17100

CGTCGCTTCT TTTAGAAAAA ATTCCGCTTC TGACAGCGAC ACACGGTGTA CCTGACTGTA   17160

CATGCCTGAA CTCACCGCAT CTAAAAACAC AATCCAGCCA ATGGCAAAAG GAGTGCGCGC   17220

ACTGTTGCGC GCACAtTTCG CACAAATCAC GCACTACAGA CTCCTGcCAC CCCAACCCAC   17280

TTGTCTCAAT GAGCACCGAC AAACCCGGAT ACTTGAGAAT CTCACGTACG ACGTCACACA   17340

GTGCAGGATA CAACACCGGA TCCCCAAATA CCGAAAGCGA AATGACCGCA CGTTCTGAAA   17400

AGTCTGCAAT GCGCCGGATC AACGCACACG CTTCCTCTTT TGGCATCAGT GAAGCATTCT   17460

CCACCTGCGC AGGAAAGAT ACTGGCCGAT ACAACGAAGA AAGCGGrTAC GCACGCGTCA   17520

ACTCAAGCGC ATAGTACGCA GGAACTGTGC GCAACGCATG CTCACGTGCG CTGATAAGCT   17580

GTGCGTgAAT TTTCTGCAGT GATATCGGTA AATGCAGCAC ACTGCAAGAA CTGCGCCTTA   17640

GAACTCGTAT AAAATTCCAG ACGCAGATGC CGCACATCAA CCGGCGCAAT CATAGTTTCT   17700

```
AGATCAAAAG AATTAATGTC TGTTTTAATG CTCTCAAAAA TGAACGAATG wCaAAACAAA    17760

TATGTGCATC CTGAGTCAAA GTAGCAAGGA TGGGGAAAAG TCCTGGCGCA ACAACTGCAG    17820

CAAATAACCC CTCCGGGTAG CCATCTGCAA AACTATACTC TGCAAGATAT TCGCGATGCT    17880

GCGCGTATAA CTGGGCGCTC GCCACACTAT CAATAAAGGG CGCGTCTGCG gCAGaACAAA    17940

GACAGCCTCA GGTTCTTCAC CCAATTGCGC GTATTCCGCA CATACGCGTG CAACATGCGC    18000

GAAGAAGGCG CTCACTCGCA TGTCATCCAG AACGTTCACG CGCAGaCGGg AAAAATAGGc    18060

GTACGCAcTG CATAAAACGC GCAACCTTCG cGGCGCTCGC gCATcCGCGT ACACACACAC    18120

CTGATGGCAA CCAGGCAACG CGTAAGCAGc TGTAACAGCA CGCTCGAAAG CGCAACGCCG    18180

CCCCACCCCG GCTGTACTCT CACAGCCCCT CACCTCCACG CACGGCACGA ACCCCGGCAC    18240

TGCTTTCATA CACGACTGct CACGCGTGCC ACACAAGGGC ACAAAGGCAT AATCGCTCAG    18300

ATCAAAGGCA CAGACCACAG CAACCGTTCC CACCGCCCCT TTATCGGCAC GCATGcAGGA    18360

GAACCTTGCT CAAAACTTTT CTGAACTTTA AAAGCACAGC CCCGAAGATC ACTCAACCGA    18420

GTGAGAAAAA GAATCATCAC ATACTCAGCT CATGCAATTC ACGCCTGCGT AAACTTCTCT    18480

TGCAAAAGGC GGGCAACTAC ACGGGGACA AAAGTAGACA CATCACCACC GAAAGAAGCA     18540

ACCTCGcGTA CCATGcTGGa ACGAAGCGcA GCATaGCaGG GcTTTGCCGc CAAAAAAACT    18600

GTTTCTAAAC CAGCGTCGAG CGCACGATGA ACCCATGCAA GATCAAACTC CTGACAGAAA    18660

TCAGTAGCAT TTCTCACACC GCGAACCAGC ACACGCGCAC CAACATCTCG AGCGTACGTA    18720

ACCACAAGCG AACGCCAAGG AAAGACGTAC ACACCCGGAC GATCCCCAAG GACTTGCCGC    18780

ATCAAATCAA CGCGCTCACA TTCTGAAAGC AAATACCTTT TCTGAACATT GACCGCAACC    18840

AACACGTGGA CCTCTGcAAA AAGACTACGC GCGCGCAGAA CAAGATCTAA ATGCCCAAAG    18900

GTAGGCGGAT CAAAAGAACC GGCGAAAATC GCCTTCACGC ACGGCAACCC CTCACGTTGT    18960

TCAGGAACAT GCGGCGAACT ACACAAAAAA CAGGCAGCAC TATTTATATT CGCACCCACT    19020

CCGTCAACTC CTCGCGGAAA ATCGGTGTTC ACGCGAGCTC TCTACGCTTC CTAGCTACAC    19080

AACGTGCGCA CCGCASGnCG CGAACGCTTC TCCTCCTACT GTGCTCTTCC CCCACCCACT    19140

GAGATCGCAA TAGCACGCGA CTTCCCGAGC GCCGCCCATC CATTGAGCGG AAAGCGCTCA    19200

TACAGCTGCA TGTAGGCAGC TGTAGCAGCA GCCGCGCGCC CTAACGCCTC TTCCATGCGA    19260

CCAACGTTAA AGAGAGCACG GGGAACGAGT GGAAAATCCT GCACGCGTGC ACTCCTCTGA    19320

TACAGCTCAC GCGCCTCCTC GAAGCGACCA CGCTCATCCG CGCAAGACGC TGCATTAAAA    19380

TAATACACGC CGGCCACGTA GCTCCTACGA GCACCATACG CTGCACGGAC ATAGGCCTGC    19440
```

```
TGTGCCTTCT CCCACTCCTT ACGCGCAAAG AAAATGTCCG CGACACAGGC CTGTGCGTAC    19500

GCATATGCAA AGCCGTCCCG CCACGGCGAA CTGGCGCACG ACTCAAGGCG CGCCAAAAGC    19560

GCATCCTCCT TCGCACGTAT GGAAGACCCT CCCCGCCCCT CCACTTCATG ACCAGAAGCG    19620

CTCACTGTAT TATCCGGTTT ACGCAATACG TCCCACTcAC GCGCTATGsk CgTnACCTCT    19680

GCTGAAGCAC GCGCGCGTAA ACGCGTCATA ACCAGCAAAC ACCCTGCACT TAGCCCTAGC    19740

CCCCCGAGGA TAGCGACGAG CACACCCACA AGCAACCGCC GGTGCAGTTC CAAGAACCGA    19800

TCAACCCGCA CTATCCCACG CCGCTGCTCA TGCATGGATC CCTCCTCCCC TCACAGAATC    19860

ACTCAACTCG AGAACCCTCA CCGACAGgcG cGAGCaGCGa AnmCCCACAC CTGcCCAAGG    19920

GAAAGAACAG CACACCGGaA CTaCCGcAAG ATACACACGG aGGCTCGGGc ACCTaCCTcA    19980

CACGcAAAAG ACCTTGCGGA GAAGCACCCA CAACAACCGA GGAGGAGCCA CCAAAACCGC    20040

AGTCAACCCA CTGCGCCCGA AAGAACCGCG TATCACACGC GCAACACACC GACCCACTCA    20100

CTTTCCAGAC GTTTGCCTCA TCAAATCACC GAGCGTAAAC GAGCCTTCGT CCTCCCCCCG    20160

CGGGGCGGaC ATATACCGAG AAAGCTCGTC ACGCTGTACC TTTCTTTGAT AGTCTCTAAC    20220

AGAAAAAGCA ACCTTCCTGT CCTTCACGTT CATATCTACG ATCACTGCCT TGACCCGGTC    20280

CCcCACTGCG TATTTCCTTA GCGCTTCACC CGGATCCCCA TCCCGATTCT CAACCAGATG    20340

CTGCTTGCGA ACAAGCCCCT CAACGCCACC GGGAACACGC ACGAAAATCC CAAAATCCGT    20400

CACGGAAGAT ACTTCCCCCT CCACGGTAGA CCCTACCCCA TAGGCGTTCG CAAACACCTG    20460

CCACGGATTG TCGCTCAACT GCTTAACACC AAGCCGAATA CGGCGCGCTT GCGGATCACA    20520

CTCGATAACC ATACACTCGA TTTCTTTACC TACCTCAAGC TCATGGTCTG CAGGACGCGT    20580

CCGCTTAACC CAGGACAGAT CATCGACGTG CAAAAAGCCG TCTATTCCCT CTTCCATTTC    20640

AATGAAAGCA CCTGCGTTCG TAACCTTTAC GATACsGsGC GTAAAGCGcG CACCCACAGG    20700

ATAACGAGCC TCTATTTCCT CCCAAGGATT CGCCGTTACC TGCTTAAGCC CCAGAGACAC    20760

CCGTCCCGCC TGGATATCAT ACCCGAGGAT CATACACTCC ACTTCATCCC CAATTTTAAC    20820

CATGTCACTG GGTTTACTCG TTTTCTTTAC CCAGCTGAAC TCACTAATAT GCGCAAgCCC    20880

CTCGATACCC TCAGCAAGTT CAATGAACGC ACCGAAATCA GCGATTTTCG TTACACGCCC    20940

CTTGACCACA TCATTCACGC CGAACTTGTT TTCAAACTCA AGCCACGGAT CCGGCTGAAA    21000

ATGCTTCAGG GACAAATTGA TACGCTTcTC CGCCTGATCC AGGCGGATAA CCTTCAACTC    21060

AATGGTTTGT CCTTTCTTCA CAAACTcGcG CGGCCGCGCC ACGTGCCCCC AGCTCATGTC    21120

ATTCACATGC AGGAGGCCAT CGAAACCGCC CAAGTCAATG AAAGCACCAA AACTCGTAAA    21180
```

GCTCTTAACC ACTCCGGATA CGGAATCTTC AATATGAACC GAATTGAAGA ACTCCTcGCG   21240

CGCCTGCCgC GCACGCTCCT CCAAATACCG GCGTCGATTA ATGACAATGT TGTCGTTGcC   21300

GCGATGCTGT TTGCTTTGGG ATATACGCTC GATATAGAAC TTAGACGTAA GCCCAATGAG   21360

ACTCTCAGGC GCGTCGACTT TCTGACAGTC CGACTGGCTG ATAGGTAAAA AgGCCATCAT   21420

CCCCGCACCC AAGTCCACTT CAAAACCACT CTTCTTTTCC GTTAGACGGA CGATCCTCCC   21480

CTCAACCGGA GTCCCGTCTC GCTCCGCATC ACGTAACTTA ACTTTCAAAC CCAAGCGATC   21540

GGCCTTCGTC TTGGAAAGCT CAGGGCCATA AGGCGTCACG CGCTCCACAT ACACCCGAAC   21600

GCCATCCCCT GcCTTCGGCG GcGCCTCAAA CTCTTCCACT GGAACGCGCC CTTCAGATTT   21660

TCCCCCGATG TCTACAAACA CCGTCCCCGC ATTAACCTGa ACCACCGTCC CCATCCTAAC   21720

AGAACCAGGT TCCGGAGCCT CAAACGAATA CCGCTCCTGc AGCTGcCGCG GcACCAATGG   21780

TGTAcCCTTC CCCTcCTGaT TTTCCACTGa ACGCTCTCCT CCCCACAAAG CTyTGCGGTG   21840

GCCTCGCGCG CGATTCTTTC ACAAACCTcC TcAATGGTCA AGCAAGAAGT ATCCAGTACA   21900

GCGGCATCAG GGGCACAACT GAGCCCCCCC AAGGTGnGnG CCCTGTnG   21948

(2) INFORMATION FOR SEQ ID NO: 64:

     (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 13518 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

AGTGTTCGCC CGACGGGAAA CGTATGGGGT CGGTACATCG AATGTTACCC CGCGCATGCA   60

CTAGAGAGCA ACACAACGCC CCGGGGAATG CGCATAGGAG CAGCAAGAGA AGCGGCGCTG   120

TACGAAGGGG AGTACGCATC CTGGTAAGGA ACAAATACGG TATAGCCACG TTCGGCAAGC   180

ATCAGTTCCA GCGTTTCAGG GTGTGCCCCG GTGCGCAGAC TGCGGGGTAT TTTTTAGAAC   240

AACCACGTTC ACGCCGACTG GGTTTTGTTT CGGGTGCGCG CGCGACTGCG GkTTTTCGGA   300

AGGkTkTGGC GCAGACGGkT TATCCAGATG AGATGTTTCG TAGGGATTGG CCAGGGAGAG   360

CAGGAsCCcT GCGTTCACCT TCGTGTTACy TAsCCGAAGG GAAAAGGAAC TGCGCAyTAC   420

GCTACGGCTG GGCCGATAAC CGGGCTCCGG GGCACAGAGG CAGACAATCA CAAGCGTGAA   480

AACACCTGCA ACGAGGAGCA GGACCCGCGC AATACGTGCT CCTGCACTGT ACAGGTCATT   540

GGGAATTCCC TGGCTGAAGG CGACAAGGCG CGGCAACTCC GTGAGACACA CCACTGAACT   600

TGAGACTAGA CTCAAGGTAA GGTCCAACGT CAGACCCTGT CCGAGGACTG TCAGTGCACC   660

GACACTCAGT GCAAGCACTG GCAGCAGGGG GATAGCGCTT ATCTTTCTGC TTTGTCGGCT   720

GAAGGGGCGA AGAAGCGCCG GCAGACTCAA ACCGAGGATG AGTAGCTCGC TGACAAGCAA   780

ATCTGACACG GCCTCCGTCT CGGTTGACTC GGTTGTACGC TTTACCAAGT TTTTCTGAGT   840

GAGmaCCACC TGTTTCCTTG TTGCGCAAGG GAACAGGTGG TGGTAGGTTT GCGCGGTGGG   900

TACCTTGTAC GTGGTAGCAA CGCCGATTGG AAACTTGGCA GACATCACCC TCCGTGCCTT   960

AGATGTATTG CGAACGGTGG ATGTAGTTGC CTGTGAAGAC ACGCGTAGGA CGCGTGCGCT   1020

CCTGTCTCAT TTTGGGATCC ATAAGCGTCT TGTTTCCTGT CGTGCACACA ATGAGGCGCA   1080

GGCGGCGCGT CGACTCATCC ATTTTTTGAG CACCCCTATT TCTGCTTTTC TCTCTCCAGA   1140

GAAGGGGAGG GGCAGGCAGA GCGCGCGGCG CACGCGTGCA CGTCCGGGTG AGACGGTAGG   1200

GACAGCTGCG CTGCAgCTcG CTGCAGAAGC AACGGGGGAA CAGGAAGTGT GTGGATCGCC   1260

GCACGCACAG GTAGCCTATG TTAGCGATGC AGGTACGCCG GGGGTCAGTG ATCCGGGAGC   1320

GGTTTTAGTG CGCGCGGTGC GGGATGCTGG GCACACGGTG GTACCGATTC CCGGTGCTTC   1380

TGCACTGACT ACTTTGCTGA GTGTTGCAGG CGTGCGAGAC AAGACCGTGC TATTCGAGGG   1440

GTTCCTTTCA CCTCACCCGG GTCGTAGGCG TGCGCGCCTG GTGCAATTGT GCGCGCAgcg   1500

TGTaGCTTTT GTTCTGTACG AGAGTCCCTA CCGGGTTCAA AAGCTTCTAG AGGATCTGGT   1560

GGCGGTGGCG CCGGAGTCGC AGGTGGTGCT GGGTCGGGAA TTGACCAAGG TGCATGAGGA   1620

GCTCTGTGTG GGGACTGCCT TGCGCGTCAT GGAGAGCTTC TGtGCcGGAC GCGytGCGGG   1680

GGGAATGCGT GTTGCTGGTT TCTGCAGAAA AATTTTAGAT CTTTATTTTT CTTACAAATT   1740

TCCGATAATG GGGCGGGGGT GGGGCTCTTn TGATGATCGA TAAGCTAATn GACTTGATCC   1800

GGTTCAGAAC CTTCGCGCCT CTTGTGCGTC TGAGCATGTG GCGCGTGCTC CAGCCGGCGA   1860

TGAGATTACT GTCTCTGCGG AAgCCCAGAA AAAGGCTGAG TTGTACTTGG CCCTGGAGGC   1920

GGTACGTTCT GCGCCTGATG TGCGTGAATA CAAAATAGCA GCTGCGGAgC agAaGCTTGC   1980

AGACCnTGCG TATCTGGAGC GGGCGCTGTc CCACGTGGTG GAGCGCTTcC TGGAGGAGCA   2040

GAATTTATAA GcCTGTAGGC AGGCTTTTTA GGTCCGGGTG AGGGCGTACG GGCTGTTGTG   2100

TTTATACCCT CAGGCGGACG CTCTCGATGT CTGGGCTGAA CAGTTCTCGC ACGTCTGAGA   2160

TACCGAGCGC CAGGAGCGCC ATGCGGTCTA CTCCTAGTCC CCAGGCCATG ACGGGGACGT   2220

GCACGCCGAG CGGGTCGGTC ACTTCTGGGC GCAGGAGACC TGCTCCTCCC AGTTCGAACC   2280

AGCCGAGTGC GGGGTGGAGT GCGTGTAGCT CGATAGAGGG CTCCGTGAAC GGAAAGTACC   2340

```
CGCCCACGTA GCGTACCTCC TGTGCACCGG CGATTTCTGT TGCGAGGATT TTGAGCATGC    2400

CTAGAAGGGT ACAGACGTTC ACATCCGTAC CCAGTACGAT TCCTTCTGTT TGGTAGAAAT    2460

CTGCAAGGTG CGTCGCATCC ACTTGGTCGT GACGGAAGCA GCGTGCGATC CCAAAATACT    2520

TACCCGGTAT GTGGGCGGTT GGCAGGTGGC GCGCTGAAAG TGCTGTTCCT TGGCTGCGTA    2580

ATAGCAGTCG GCGGGTAAAA TCCCGATCGA ACGAGTAGCG CCAGCCGAGG CTCCCGCTAT    2640

CTGctCCGCG TTCATGCGTC GCCGCAACGC GGGAGAGGAA CGGCTCTGGG ATTGTAGGTG    2700

CGTGCGTTGG GTGTTTGAGG TAATACACAT CATGAATGTC CCGTGCAGGA TGGAACTGTG    2760

GCATGAACAG CGCGTCCGCG TTCCAGAAGT CTGTTTCCAC CAGTGGCCCG TCAAATTCCT    2820

GAAAGCCAAG TGCCACCAGA CGATCTTTGA TGTGTTCGAG GAAATCTGCG TAGGCATTAG    2880

ATCGGCCGGG GATGATGCGG GCAGGGGGAA TGTGAACGTT GTAGCGGCGT AGATGTTGTG    2940

TCTTCCACGC GCCACTTTTT AGGCACTCGA CGGTGAGTGC ACCTATCTCG TTTCCGGTAA    3000

GCCcTGCGGT ATGCAGCGCC TCCTGCACGG CACGGGCGGT GGGGGTAAAG GTGAACGTTA    3060

CCCTCTCGCG TACACTGACT TTGAAGAGGC TGTCGctTGC GCCCCGCTTT TTTGCTATGC    3120

GTTCCATTAC ACGTCGCTCA TCATCAGAGA GCTCAGATTC AAAGAGGGTG CCTGGAGGAG    3180

TGTCCGAGGC CTCTGACGGG GAAGCGACGC GTGCAGCGGC GCGCTGAAGC AAGGTGCGCG    3240

TGAGGGACAT GCGAaTCACT TACGTGCGGT GAAACGATGT GTATGCGTTT TTCACCGTCC    3300

ATACGGAGGA TACCCTCCTG CGCTAGGATA CCGAACGCTG AACCTACATC CTTTGGTGCG    3360

AGCGTGAsCG CATGGGCAAG CTCAGGGAGA CTGAGCCCGT TGCAAAGTGG GGGGCGAGGG    3420

TGTAGGTGCT CGGCTGCATC TGCAATAGCG GTAAGGGAAG GGGGGGAAGA CAAGAAGGTG    3480

AGCATACGCT CCTCTGCAGT ACCGTCGCTA GCGGCAGCAT AGCCGCAGGG GGTGAGTTCA    3540

AAGGACCGCA TCTGTTCCCg CTGGTGCTCT TCGATGATTC GCTTTGCCCG AAGCCAGGAA    3600

AACGCTTGGT TTGCGTGTCC TTCCTTAAAG CCTAGCCGGG AGATGAGCAA CGAAGTCGAA    3660

AGGATCTCAT CCATTGCGCA GTTCTTGAGG ACTTTGATCT CAAGGGGATG CAGCTTGTGC    3720

ACAAGCGTGT TCAGATCGGC TTTACCTGTC ATGCGCGGCA TCATAACGTA TTTTCGCGCC    3780

GTTTAGATGT AGGCTGTTTC TCAATTTTTC GTTCTGCATC AGGGTACATC TGCCGTAGTG    3840

TGGAGAAATA GATGGAACGA TCCTGGGAGG GTATTGACAG AGGTGTGCCC GTCTATGGTA    3900

TGGGTTGCAC CATGGTTGCT CCTGGGATGC CCGCAGCTTT TGCTCGCTCG GTGGTTCGCG    3960

CGTCTGCGTG GGTGGGTGTG GCGCTTATGT GCGTTGCGTG GTCGGTCTCC GCCGCGGAGG    4020

GCACGCGGTC GGGTGGGCAG GCTCAGGAAC GCTTAAGTTC CTGGCGCCAG GTTGTGCAGC    4080
```

```
GCATGGAGGT ACATCTACGT GCGGCGTACA CCTTTTTTGA GAGTGGGGAT AGTGATCGCG    4140

CCTATGAGCA GATAGATAAG GCGTACTTTC GCTACTATGA GGCGAAGGGC ATGGAGAAGA    4200

TCACCATGGG GTATCTGTCC GGTGCGCGTA AGGCGGCGG GGAGAACGCG TTTTTCGCGT    4260

ATCGGCGTTC GGTGCGGGGT GCGCGTGATT TGGCGGGCGT TGCCTTCTGC AGGGACAAGC    4320

TGGTTACCAT GTTGTATGAG GACGCGCGTG CGCTGGATGG GGTTGCGCGT GGTCGGGCGG    4380

GCTTTGCGGC GCATATCGCC ACGTTTGTTG CCTCGTGCGT GTTGGTGCTG CGCGAGGGAA    4440

TTGAGGCAAT TTTGGTTATC GCAGCGATTG TTGCGTATCT GGTGAAGACT GGTAAGGAGC    4500

GGTGCTGCGC TGCGGTGTAT GCGGGAGCGG GCGCGGGTGT TCTGTTCAGT GTCGTGCTTG    4560

CGGTGATGAT AGTCCGGGTG TTGGGTTCGG AAGGTGGTGC GGCGCAGGAG ATTATCGAGG    4620

GTGTTGGTAT GTTCTTCGCA GCGGCGATGC TCTTTTACGT GAGTAACTGG ATGTTGTCCA    4680

AGGCGAGGGC ATGTGCTTGG GATCGCTATA TCCGTCAGAA AGTTGAGCGG TCGGTGTCTC    4740

GGGGTAATCA GTGGGCGCTC GTGGCCACTG CCTTCCTCGC AGTGGCGCGG GAAGGGGCGG    4800

AGCTTATTCT TTTCTTTCGA GGCATCCCAG TTGCGGGGCC ATATGGGCGG CTGGCTGTGT    4860

GGGCAGCGGT TACTGTTTCT GCCTTGGTTC TGGTGGGTGT GTTCGTGGCG ATCCGTTTTC    4920

TGTCAGTGCG ACTTCCGTTG AGGCCTTTTT TTGTTGCCAC GGGCGCGGTG ATGTACTTGC    4980

TATGTTTCTC TTTCGTGGGT AAGGGTGTCA GCGAGCTGCA GGAGGCAGGT GTGGTCAGTC    5040

GAAGTACGGC ACCGTGGATG CATGGGTGGA GTTTTGATTT TCTGGGCATC TACCCGACCT    5100

ATGAGGGTCT GGCCCCTCAA GCGTTTGTGG TGGCGTTGGT GGTGCTTTCG GCGGTATGGT    5160

GGTGTGGTGG TCTCTGCCGT GGCGCATCCA GCACGTAGGC TTGGGACGGC TGTGTCGCGT    5220

CCTACTGGGG CCGGGTGTGT GCTGCGCCGT GGAGATTTCC ATTTGTTTTT CTATAATGGT    5280

GAGGAAAAGA AGCGCTGGAC GGGAGAAGGC GTTTTGAAAA GGAGGGGCGC GTGACGCCCC    5340

AGGGGAGTGA AGAATGAAGA GGGTGAGTTT GCTCGGGAGC GCACCATTTT TGCGTTGGTT    5400

TTTTCCGCGT GCGGGGcGGT GGAGAGCATC AGCACGGTGA GGAGATGATG GCCGCCGTTC    5460

CTGCTCCAGA TGCAGAGGGG GCGGCCGGTT TTGATGAGTT TCCTATAGGC GAGGATCGGG    5520

ATGTGGGGCC CTTGCATGTG GGArGGGTGT ATTTTCAGCC GGTTGAGATG CATCCGGCTC    5580

CAGGAGCACA GCCGTCGAAG GAAGAGGCGG ACTGTCACAT AGAAGCGGAT ATCCACGCAA    5640

ATGAGGCGGG TAAAGATTTA GGGTATGGAG TCGGGGATTT TGTGCCGTAT CTCCGAGTTG    5700

TTGCTTTCCT CCAGAAGCAT GGCTCTGAGA AGGTGCAAAA GGTGATGTTT GCGCCCATGA    5760

ACGCAGGGaC GGTCCGCATT ATGGGGCGAA CGTGAAGTTT GAAGAGGGGC TTGGTACGTA    5820
```

```
CAAGGTACGT TTCGAGATCG CTGCACCCTC GCATGATGAG TACTCGCTAC ATATTGATGA          5880

GCAAACTGGG GTTTCCGGAA GGTTCTGGAG CGAGCCATTA GTTGCAGAGT GGGATGATTT          5940

TGAATGGAAG GGGCCTCAGT GGTAGGGACG TTCAGAAGGT CCGAGGGTGC GCGCGCATAA          6000

GGGCGTTCTT TGTTCAGTAA GACAGGCGGG TAGTGCAGTG CGTGGCGCTG CTCGCCGGGT          6060

CCGTTTTGAG GGTGTGGGTT TTGACACGCA gTTATTTTTT TGAAAGTTCT CCTGCGCGTT          6120

CTTCTGTCTC CGTGGGGTTG TGCGGTGTAC AGAACGGGGG GGGGTGTCGT GAGTGCGGGT          6180

ATGAAAGTCT TGGTGTACGC GGTGGCGCTG GGGTTCGGGT GCGGGGGTGT GGTGCACATG          6240

CGGGAGGGGG ACACCTACCA ACAACTCCTC GAGCACCGCA TTGCAAATGG TCGGGAGTTT          6300

TCGCGGGTGT TTGCgCAGGC ACAGGTTGAC GAAGCTGAGC ACAATGAAGT TCGGACAAAG          6360

ACGGCGGGAA GTGTGCAAAT TGGCACGGGA GACGTGCTCT TCAACAAGAA GAATGGCAAT          6420

GGTGCTAACG GCTACAAGGT GGAGATGGCG CCGCATTTGA GTATTGCGTC CCCCTTTATA          6480

GGAAATTCTC GGCTGAATCT TGTTGCCCCC CGCAAGCTTG ACGGTGTCAC AAGTACCTCC          6540

ACCGTGTCGG TGGATTACAC TACCGATTTT TACTCCTCCG TTCGTCCAAC ATACCTGAAC          6600

TCCCTCAAGG AAAAGACATA TCAGAAGGAG AAGAGCGGTT CGGCGCTGCG TGATGGGCGC          6660

AGGCTAGTGG AACGGGAGTT TTTGCAGGAA GTACAGCGTC TGTACGGTAG TTACGCGGAC          6720

CAGGTGCGCG CAAGTTTGGA GTTGGTGCGC GCGCGGTTGC GTTTTGAGTC AGTAAAGAGA          6780

CAGGGATATC AAGAGGATTC GGCGTATTTT CAGAGCGCAC AGCTTGCACA GGTGCGGGCG          6840

GAACGCGCCC GGGCACAGGC CAGGCAGCGC TTTGACCTTG AGTACACGCG GTTTGCAGCG          6900

CGCAACGGGG TGGCCTACGA GGACGATGAG CGCGACGGTT TTTTACGCGA TTTGGCGGTT          6960

GCAgTGCCGC TTGAGCCGGC GATGGCGGTG ACTCAGTGcG nCAGGGGAGC GGGGGCGcGA          7020

GTATTGTGAT GCGCAGGAcC GCTGCGAGCG CGTCATTGCC CAGCGAGGTA CAGATTACTC          7080

CCCCTTTCGC ACAAGCGCGC GCGTGTACTT TACCGATGGG GAAGAAAACA AGCAATTAAC          7140

TAACGGCATG GCGCCAGCTG CTCCTAGCAC TACGAGCACG TATGGGGGCA CGTTCAACAT          7200

GGCGTTTCCC GGCGGGGATT CCAGTTTTAC CGTGCAGAAT AGTAAAGGGC TGGCGGGGAT          7260

CCTAGcGAAT TTTGAGTGGA GCCCGATACG CACGCcTATC GCTCGCTGGA CTACACGGCA          7320

GAGCGCGCAG AGCGTGTcTT TGACGAAGTT GAGCTTCAGG CAAAGGGTGA TCGGTCGAAT          7380

AAGTTGTTCG GTGCTATAGA CGCGCAGGGG GACTCAGTGC TGGTgcTGCG CGGGGTGGAT          7440

TTGCAGACAC TGGATAACGC GCGCAAGAAG GCACGGTTGC AGAAAGAACG CTTGGAACGT          7500

GGTATCATCG GAAGGCTGGA GTACGAGGCA GCGCGTTCGG AGTATCTTCT GGCGCTTGCT          7560
```

```
TCTGTGGCAG AGGCAAAGGC GCGGGCGATT ATTTTTAACA CCGACCTTGC GTGTGCCTAC    7620

GGGGTGGGTG CGGACGCCGC CGCCGCTCAG TTGACCCAAG AGGAAATGGT GGTCTCTGAG    7680

AAAAAGGATG CTGAAGAAAA GAAAGAGAGG TCTTCGTGAG CGTAAgTTAT CGTGGCCCGA    7740

GGTGGTCTTC GTTCGTCCAC GTGTCGCAGC ATTCGTGTAG GTTCGTAGCT CCTACGTGCg    7800

CTGAGGGTGC TCAGGGGTGC TCTGAGTTTG GGGCGTTCCC TGTTTTTGAG GAAAGGGGAA    7860

TGTGCGCGGC GCGGCGTATG CGCAGGGCGG CAATTGCCGC GTGCTGTGTG. TTCGCGCGCG    7920

GTGCGGCGGC CAATCCGTAC CAGCAGCTAT TGCGCCACCG CCTGGAAGCG TTGCGGCCGG    7980

GTGCCCGCGC GCAAATAGAG TTTGATGTGG CGCACTGTGG GTATGAGAAG GCgCGtTGCG    8040

CTcAGCAGGT ACGTACGTGT TGGGCAGTGA GCTTGAAATC AGAGGACACT CgGCGGGGGA    8100

TTTTGGGCTC CCTCGCTTTG GAATAAAGCC CATTATCGGC GTGAGAAGTC CGCGCTACAA    8160

TAACCTGGTC GTGTCCATCG ACACCGCAAG GtAACTAGCA TAGGGAATAT ATCCCGGATA    8220

AACGCGGATA TAGGGGTGGA TTTGTATTCT AACGTGCGGG GGCGCGAgcT CATTCGTATG    8280

CGTCGTGCAG AGCAaAAGaA AAGGCGGCGC AGAACGGTGA ACGAATTAAA TCGCCGTcGG    8340

TGGAGCTAGC GCTCATCGAT GAGCTGGAAG TGCTTTTTAC CCGCGCGCAG TCGCTCGTGC    8400

GGCGAGAGTT TCATATGGGG GATGCGCGTT TGGTGCACCT GCGCACGCGT GCGsCAGGTT    8460

TTTCTGAGCA CTCTGAAAAG GCCCGGCGCG TCCGTTTGGC GTACGACCGC ACACAGCGTG    8520

AGTTTGAACA AGAAGAGCGC CTGTTTGCGC AGGTGTGTGA TCCCTTCGCT GCCGTCTGCG    8580

CAGTGGGCGG AGGGGATGAA GCGCGGAGAG ACTTTTTGCT GCAGCTTGCA GAGGCGGTGC    8640

CGCGCGAGGT ACCGCTCTCG CTCGTTTCCT TGCATGCTAC AGATGCGCAC AGCCTTGCGG    8700

CGGcGCAgGA GATGGCACTG CTTGAACGCG CCGCGCAGAT tCGGAGCGTG ATTTGTACGC    8760

TGTGCGGTGG GCGCTGTTGT GAGCATGGGT ACACGGAAGA CTTTCATTCT GTTTAAGGGA    8820

GACGGAACCG AGTCGCTTGA AGGTTCCGGG ACGGTGGCGC TGCATATGCC CAGCGTGAAC    8880

GCGCAGGTAG AGGTAAAGGT GCCCTACGCG GAGAGGGGTA AGCATTCCCG TGACAAGGTG    8940

GGAGTGTACG GGAAGTCGCA GTGGAATCCG CTTGAAATTG CCTATAAGGT GTTCGAAAGA    9000

CGGGAGGAGC GGGCGCAAGA GCAAGAACAG GAGCAGTATT GTGAAGATTC CCTGGCGCGT    9060

GAAAcGcgGA aGATGGAGGG GTTAGAGGTG CAGGGCAAAC AGCTTTTTGC AGCACAAGAA    9120

ACCGCCTTGc GCACGCGCGA GGCgCTGCGT TTAGATCTTG CCaAGGTGGa rCgCGCCcgG    9180

CGCGCGGGkT AGTGGGAGGA AATCGCCTCG CGCGTGCGCG kTGTGAGTAT GCCGTGGCGC    9240

AgcTGCGTGC GGCGTGCGCG AAGTTGCATA TGTTGCGTTT TAATCTGGGA GTGGTACGCG    9300
```

```
CATTTGGCCT GGTGCCACAG GTGGCGCCGT GAGCGGTCCG CGGGGTGGTT CGTATGGCAA    9360

GCGCCGGGCG GCCGTGCGTG TGTTCGCCGG AAGCGTGCTG TGGCATCATG CAGTGTTGGG    9420

TGGGATGGGC gGsTGCGcTC ACCGCCAGCG AGTTAACGCC CGGCGCACCG CCGGCGGCAA    9480

GCGCCCGGGC GGCCGCGCAA GAAACGGGAA CCGACnTCTA CCAGCGCGTG GTGCGCTATC    9540

GGCTGCAGCG CAgTACGGCG GCGGCGCAGg CTGTCCGACG GCAGACGATA ACACAGAGCC    9600

AGTACGATAA GCAGCGGCTT GATTCCTTGG TGCGCCTTTC TATCGCAGCC GGGGACATTG    9660

CGTGGAACgc CGATGGGGTA AAGTTTCGCA TTACGCCCAA GGCnTCGGTG GCATTCCCTT    9720

CTTTTTATAA CCTGACCACC CATTTTGGTA TGACGGTAAC GCAGCCGAAC GGTGCCGCCG    9780

GGGGAGGAGs skGwnGAGGG GGAGGAGGCG ACTGGCAAAA GACgCTCGAC GCGGGGgCAG    9840

GCATTGATTT GTACTCGTCG GTGCGTCGCA GCCATGTGTT TGCGGTGAAC ACCAAGTACG    9900

ArGsmnTGCG TGATGCGCAA GAAGCGCTCG CCTGTGAGCC GCACGTAAGT GAGAAGCAGG    9960

TGCTCGAGGA CATGCGCCGG ATGTTGGATT CCTACGTGCA GCTGTTGCAC GCGCAGGAGT   10020

CGTTTGCGCA AAAGCAGAAC gcAGAGCGAT CAGTGCAGGT GGCTGGATAC ACGGACCGCT   10080

CCATTGTGTA mCGCGCAgCA GCGCTCGAGC GGGAGCGrGC ACAGGACGCG CTCAAGGTGG   10140

CGCAAGACGC GTTTGACGGA GAGTACCGGG ATTTTATCAT CTCTGCTGGT CAGGAATTTT   10200

TAGAAAAACG TGCGGATCAG GAGCGCTTTC TGCTCGCGCT GGCTGAAAGC GTTCCTGAAA   10260

TGCCGCTGGT GTCCACCGAG CAtGCGAGGC AGATACGTCC CGCCCTCTGc GCAACGCGCG   10320

TGAGGCAGCA GATAACGAGC GCGAGGAACG GGCGGTACAG AACTTTCCCG TGGCGCTTCG   10380

TCTTGACACC CGCTTTACCC TAGATGAAGG AACCGGGGAG CTTTCCGTTG CGTTTCCAAG   10440

CGTCAAAATA ACCAGCGCCC TGGCCATAGG TTACACCGGT ACGCTCAAAA GCATTGGCGG   10500

GTCTCTGGAC TGGCATCCGT TTGAAATCCG GTACGCGCAT TTGCGAGGAA AAAATCAGCG   10560

CCTGCACGAT GCGTTAGGGG CACGGGAGTA TGCACAGAAA AAGGAGCAGC AGGAGAAAGT   10620

AATCGCAGAC CTCCCACAGC GTGCAGAGGA TATCCTCTGG GAGCGTGAAA CTGCACGCGC   10680

AGAGCGGGAC ACGTACGCAG AAAGCGCCCG CGCGCACAGG AAAGGACTTG ATCGGGGAGT   10740

TATCGGCGCG CGTGGcTACG CGGCAGTACA TTTGGACTAC GTACGGGCGG TTATCAATTT   10800

GGCGAAGgCG AATGTAGACG CGCTCATTTT TAACATCGAC GCGCGCGTAG ATTTTCTTTC   10860

TTCTGGAACC CAAACATGAA CATGGGGAGT CTGGTATCTT ATGCTGCAAT CCAACGGGGG   10920

GTAAGGTGAT CGCACGCAGG ATGCTTTGCG CGCGCCCGTG GGGGCCGTCG TGCGTGGTGT   10980

GCGCTCTGTG TGGGGCGCTT GCCGCCTTGG TGCCAGCAGT CGGTGCGCAG GAACAGGCAG   11040
```

TGCCTGCGCC GGGGACGCCG GCTCCTCCCG CACACACGGC TTCAGAAGCG GTGCCTCCTG   11100

CGCCAGAGCC CCGTGCGGAA GGGGAGCAGC CGTCTCCTCT TGTCCCCACG cTCTGCCGGT   11160

CCCTGGAGGG GCAGTGGCTG CACGCGCAGC GCCsGGCACA GTCGGTCCGC GGCTGTGGGA   11220

GCAGCTGCTG CAGTGGCGCG TGCAGCACGG TGACGAACAC CAGGCGCCGC AAATGGCCTA   11280

CGAAATTGCC GCGAACAATT ACGACATTGC GTTGGTAAAG TCCATCGTGG ATCTGAGGAT   11340

GGGGACTGGA CACATACACC ACAACCTGAA TGGGAACGGG GCCGGGGGTA TGGCAAACGG   11400

TACGCCGACG CTTTCTCCCT ACGTGCATCT TTTTTTTCCG ACCTATCAGA ATTTGAGTTT   11460

AAAAGCGGAT ATTGCGATCA AGACCAACAC CCcTTCGGCA GACGTGACCG CGCTCTTTGG   11520

TATGGATCTG TACTCCAAGG TGCGGCGGCA GCATCAGCTG CAGGTGCGGC GTGCGCGCAA   11580

TAGCATGCTT GACGCGTTTG CGGCGCACtG CGGGGGCAGc ACgctGCGCG GGAAGCGTTC   11640

CTGGCTGAGC TCGATGAGCT GCTAAGCGCA TACAGCACGC TGCTTGAAGC ACAGGTAACC   11700

GAGCAGGAGT GCACGCGCCT AGTGCGCACG ATGCGCATAC AGCGCTACCA AGCGCATTCG   11760

GTAAAGTTGC GctCCgCAAC GCTCAAGCAC GCACGCGCAG AGAGAGTTGC CCGTCGTGCG   11820

CGCAAGACGT TCACCGCCCT GTATCAGGAT TTTGTGCGCA AGTGCGGGGC CTTTGAAGGA   11880

AATGATCCGG AAACATTCAT GCTCCATCTT GCGCAGGTAG TTCCGCAGGA GCCCGTATCT   11940

TCTAnCCGCA CTGCTTTCAG TGGAAAATGA CTGGGAGTTT CTTAAGAACA GGGAAGATTT   12000

GGAAACTCAG GCTGAAGCGC GTGCAGTGGA TGCTATCTCG TACGGGTTTA ATGTGGAGTC   12060

TGGGGTGGGG TCTGAGGGTA AGTCATTGAA gAGAATATTG GCAAATGTCA GAAtGGACTT   12120

TCCCGGCGGT GGCTTTTGGC TTGGATTGAA CTTACCGTAC CCGcAGTGGT CCCGTGTGGA   12180

GGTAAAATTT CGGCTCACGT GGGACCCGCT TTCCATTAAG TATcAGGAGC TTTCACGGCA   12240

GACACTGCAG CTTCATGAGC GGCTCAGTGC GCTTAAGCTT CAAGACGCGT ACGAAGCTTC   12300

TGAGCGTAAG GTGCTTGGCC TGCGCCACAC CGCCGAGTCG CTCGGCTGGG AACAAGAGGC   12360

GGCACTCACC GAACTGAATA TTCTCAGGCG GAGTGCGCAA ACGCACCAGA AGTGGCTGGA   12420

AAGAGGAGCT ATCGGCGCGC ATCAGCACGC CCGGGCCCAG CACGCGTACC TACAGGCGCT   12480

CATCACGTTG GCCAAGATCA ACATTAAAAT ACTAAAGTTT AACCTTGAAA CTGCGTCTTC   12540

GTTCAGACCA GTACTCTAAA GAATACCCCA AGAAGGAAGT TGTATGACCA CAGCACAGAA   12600

ACTCCTACAC AGAAAATCGA CCATCGCCAT GGTGGTCGGA ATTCTCGCCT TCTTATTTGT   12660

TCTTCCCCGC TTGGTGCGGG CGCTGCGTCG GGTTCCGCCG CCTACCCTCA GTGTGAGTAA   12720

GGAGGTGGTG CTCAATAGGA TTGAGATTTC GGGGTACATC GAAGCGGCTC AGCACCAAAA   12780

GCTTGAGTCC CCTGGTGAGG GAATCGTGCG CACCGTACGG GTGCAAGAGG GAGATACGGT   12840

GAAGAAGGGG CAACTCCTCT TTTCGCTTGA AAACTCTCAC CAGCAGCTTG ACCTTGCCGA   12900

GCATGAGTTT GCAATCGAAC AAGAAGAAAT TAACGGTGTT TCTAAAAAAA TGGAGATCAT   12960

GAAGCTAAAG AGAAATATGC TCCAAAAAAG ACTGAGGGAA CGCTACGTCA CTGCCCAGTT   13020

TGATGGCGTT GTTGCCGCTT TTAAGCTCTC TCCCGGACAG TACGCGAAAC CTCAAGATTA   13080

CTTTGGCACT CTCATCGATC GCTCTTACTT CAAGGCAAAT GTCGAGATTC CTGAGGTGGA   13140

CGCTTCGCGC CTCAAGGTAG GGCAGCGCGT TGAAATTTCT TTTCCCGCAG AACCAAGCGT   13200

GAAAGCGGTG GGGAGTGTCA CTTCCTATCC GTCCATCGCG CGCGTTACCA GTGTCGGGCG   13260

CACCGTGGTT GACGCCTCCA TCAGGATCGA TGAATTGCCA GAAATACTGC CGGGTTATTC   13320

CTTCAGCGGG GCAATTGTTG CCGGGGAGCA gGAGGAAATT TTAGTCCTGA AAGCCAAGAC   13380

GGnCTCCGGT ACGAAGAAGG GTGCTCCGTT CGTGGAnCGA GTGCTCCCCA GCGGTAAGAT   13440

AAAGTCTGTG GCCGGTTACG GTGGAGCCGT ATGTTnCCTG GCTTTGGTCA AAAATAAATT   13500

TCTGGGGCTG GGGGGCGG   13518

(2) INFORMATION FOR SEQ ID NO: 65:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 4448 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65:

AAAATGACAn AAGCACAACG GGnAGCGGTT GGAGGTTGCC GGTGACATGC AGCGCATGAT   60

GAATGCAGGG CGCGCGAAAC AACGCACGGC GCACAGGAAG CGCGTGAAAG TGTCACGCCC   120

TGCTACGGCC TTAAGGTGGT GGATGCCCAG CACTTGTTAT CGGAAATCGT GCTCGTTGAT   180

CCAAAGACTG AAGCGCCTTT GCGTTCTTCT TCCCTACGGA CCGTCCGCAA CCGGCTCCTG   240

TACAGCGAGC CTCACGCGCT CGTCGCCATT GCTGACACGA CAGGGAACGG CACCGTCCGC   300

CTCGTGCACA TAGACCCAAA GACGCTGGAG GTAACCAAAG AGAGTACCCA GCGTATAGTG   360

CGCAAAGTTT TCTCTTGAGG GAAGAGGAGC AtACTATGCG GTGATCGACG AAAATGGCAG   420

CCACTTCCTG GGACGCTTTA CCAAAAATCT TGAGCTGACT ACTCGTTCTG CAGCGCnGnT   480

GACGCCTTAT ACCGCCGTCA CCGTCACTCC GCGCGGAATT ATGGTGCAAA CAAAAGAGAA   540

AGGTATGGCC CTATTGCACA CACGGACGCT CGCCGACGCG CTACCCAGAA CATGAGCAGA   600

```
AAGAACGCGA ATCAGAAAAT TGTGTGGCCG CTTACGACGA AACTCATCGG TATTATCAGC   660

ACGGTGGTGG TGCTCGCCAC CATTGCGGTT ACGGCCATGG CTTCCTGGTT CTTCGCTTcG   720

AGTCTGCACA GTAATGCTGA GCTTAATAAT CTTGCCGCTG CGGAGAACTT TGCTGCGCAA   780

ATCAAAGGGG AATTCGAAGC CATTGCCACA AGCGCCAAGT CCTTCGTTTC CCTTGGCCTC   840

AGAAgTGGCG CGCGCATGCA CTCGCGGTCC GCACTTTCTA AAGATTTTTT TTCCTTTTAC   900

CCGCGCATCG GCTACATCGG AGTCGGCGGT GTAGCCGAGC TGTGGAACGG TGACTTCTTC   960

AAGAAAAATC AGCTGCGGGT GGCAGACGCT CGTCGCTTCC TAGCTGACAA CGCACAGGTT  1020

ATTTCCACAC TTCAAACTGC CCCAGCCACG CTCAACGCCG CCCCCTGGTT TAAAGCGCAG  1080

ATCATTGCTA TCGTCGCGCC CTTTGAAGTT GACGGcGCTA CGCGTAACGT TGTGGTTATC  1140

TTCTCAGCGG ATGTCGTTCA GCACCTGCTA GAATCTGGAG CCTCCTCCGG AACCATGTAT  1200

GCCGTCACCT GGGCGGGGAA CTCCCTGTAC CACCCGGAAT ACTCTCTCAA TtACAGCAAC  1260

ATtAACTTGC AGGACTCGCC CGTTGTGCGC GATTTACGCG AATCTACACA GCTGACCAAA  1320

CAAATCAGCT TCATCGGCAC GGACAACAaG CGCTACTTCG GCGCGTTCGC CAaGCAAACC  1380

TTTGGAAAGT TCGCCmTAGT CCTAGAAACG CCTATGAGTG TGGTGTACCA GGCAGTATAT  1440

TACGCGATTA TCCTCGACGG TATCCTCACC GGCATGGTGC TCCTCGCCTC TATCTTGCTT  1500

GTCTGGTTCA TTGCGCAGTC TATCACCCGC CCTATCCTTA CCCTCGTCGG CGCAACGCAC  1560

GCTATCAGCT CAGGACAGTT CCTCCTGGAT ATCAAGCCTT CAAGCAAAGA CGAAATTGGC  1620

CTCCTCACCG AAACATTCGT GAGTATGGGG CGTGGTCTGG CAGAACGGGA ACGCATGAAA  1680

GAAGCGTTTG GCAAATTTGT AAATAGAGAC ATCGCAGAGA AGGCCATGAA GGGAGAGCTC  1740

GCACTGGGAG GGGAACGGAA AACCGCTACC ATTTTTTTCT CAGACGTGCG CTCCTTTACT  1800

GAGATGTCGG AGAAGCTTCC CCCTGAGGAC GTAtAGAGTT TCTCAACGAG TACATGAGCT  1860

GTATGGTAGA CTGCATCGAG CAGACAGGCG GCGTGGTGGA CAAGTTTATT GGAGATGCGA  1920

TTATGGCGAT ATGGGGAGCG CCAGTTTCCC TCGGCTCTGC ACGCTTAGAC GCATTGCAGA  1980

GCATGAAAGC GGTCTTCCTC ATGCGCGAAA GCCTTATTCA ACTGAACGAA AAGCGCGTCG  2040

CATGCTCAAA GCCTCGCATT GGCATCGGAT GCGGCGTAAA CACAGGCTCC TGCGTCGCAG  2100

GTCAAATCGG CTCTTCCAAA CGTATGGAAT ACACCGTCAT CGGAGACGCG GTGAACACCG  2160

CAAGCAGGAT CGAAGCACTG AATAACCcGT TCGGCACTGA CTTTCTTATC TCCGAAAACA  2220

CATATGAGCT TGTTAAAGAT ATGCTTATAG TGGAGAAAAT GCCCCCCATA ACGGTAAAAG  2280

GAAAACGAGA ACCACTGAAT GTGTACGCTG CTATCAATCT AAAGGGGCAT GACGGACCGC  2340
```

```
AGACGCTCGA TGAGCTGCGT GCACTTCTTT CCATTGAAAA GCCGGGGCTT TCTGCCGACC   2400

CTGACTTCGA AGAAAAGAAG TGTGAAGTTA TCTAAGCAGG ATGCCACGGT TACGGTCGTT   2460

ATTCTCCTCC TTATCCTGCT TCTCGGCTGG GGCTACTCCC GCGCGCTCCG TCTGTCCCAG   2520

GGGAAGGGAA ATCCAATCGG ACGGGTTTTT TTTTATAAAA AAACCGCAAC CCGCAAAAAA   2580

AACAACCAAG CCTTATGGCT CAAACTCAAA GACGGGGTGC CCGTCTACCA TCGGnGyAss   2640

TGCGCACCAC CACCGGTTCT GAAGCTGTCA TTGTGTTCAC TGATAACAGC AGGCTCGACA   2700

TTGCAGAAAA TACCATGGTG CGCATCAGTc ACACAGGAAT GAAAAGAAG GATGTACGTT   2760

TGGTCACAGG AGCGATTACG tACGCaCGCG CCGCTGGGAA TCCAGCAGCG CATACCGTAC   2820

ATGTAGGAAA GACAACCATC TCGCTTTCTG GAGACGGTCA GGTGAATGTG CGCGGAGGCG   2880

AACGCGATTC AAcTGTCGAG ATAGCACGCG GTGAGGCACT CCTTCACGAT GCGCAGGGAC   2940

AGACAcTTCC CCTTCAGACG TTCACCCAAC TTGCTACTTC CCGGGAGGAT GGCACTGTGC   3000

GCATTCTGCA CCCCACCTTT GTCCCTCTCC TACCCGACCA AGATGCACTT CTCCTGACTG   3060

CCGAGCACAC CAGATCTGTG GGCTTTGTCT GGCTCGGCGA TGCCACGACG GTACAGCCGA   3120

GCGTCCGTCT CCAAATTAGC CGATACGCGG ACTTCTCGGT TATTGAAACG GAAAGAAAAC   3180

TTACCCTTCC GCATGAGGCA AACGCCTCGA GGACAACATT CAAAACCAGC GAACGACTCG   3240

GGGAAGGACG CTGGTTTTGG CGCCTGGTCC CGCAGAACGG CACGcGTCAg CGCCCCGTTC   3300

CTTTTCTGTG CGTCGCGCGC GtAAGGTGAT GCTGcACACG CCGCGTGCTC AGGCAGTACT   3360

CTCCTATCGG GATGCGATTC CTCCTACCCT TTTTTCCTGG ACGTCTGTAG AAGACGTGGA   3420

ACAGTACCGG CTACTGCTTT CTTCCCGGGC CGACTTTAGC GCGGATGTGA AGACATTCTC   3480

TTTGCGTACG CCGGAGATCT CGGTACCCGG GCTCGGCGAG GGAACGTATT TCTGGAAGgT   3540

AGTACCTCGC TTTGATGAGG GAATAGAAGA CCCAGTCTTT GCTTCTGAGG TAGGAACCTT   3600

CTCCATCAAA CAGGGAAAGG AGCTGCATGC GCCCGTTGCG CTCTTTCCCG CCGAGGACGA   3660

GGTGCTCGAA CACGCGGATC GGGAAAATCG CATGGTAATC TTTACCTGCG AGCCAATACC   3720

AGAAGCACGG CGCTATGTCT GGACGGTTAA AAACATGGAT GCAAACGCGT CCCCGCTTGT   3780

GACTACCACG TCGGTACCCT TTCTTACCGT TCCCATGCGG AGCCTGCGTG CACGATTGCA   3840

GGAAGGAACA TATCAGTGGC AGGTAGCGTG GGAAACGCGT CGGAGCGATC GCTCCCCCTA   3900

CTCGGCACTG CGCGCGTTCA CGGTCATTGA AGGAATGCAC GCGTGGGAAG AGGAGCCAGA   3960

GACGCGTGAC TTGATTkCGC TCcGCTcCTT CCTTTGgyTG CGCGACATGC CAGCACTCAT   4020

TACTGAAAAA TACCTTTTGC AGCATCGcGC GTTGCGTTGT AAGTGGACGG CGGTGCACAA   4080
```

CGCACAGCGG TATACGGTGA CGTTAAAAAA CAAGAAGACA GATGCGGTAC TGCAAACGGC   4140

AACTACCACA GGGGTGGAGT TCTCATTTAC CAACTTAGCG CACCTTGAGG AAGGGTCATT   4200

TCATTGGGTC ATACAGGCAC ACACAGAGCA GGAAGGCTAT GAGCCTGCAA GTGCACAGGT   4260

GGTGCGCGCG TTCACCATAC GGGTGTCTGA ACTTGAAAGG CCGCGCGCAA AAGAAATTGT   4320

CCATTATGAG TATCATTAGC CGCGTGTGTA TACCGTGTGC GGTGCTGCTG TTTGCGCAAC   4380

TGCACGCGAA GGAACTCGTC CACGTATCTC AGTTAAAAGA ACAGGAAGCG CGTATCAGCT   4440

GGCAGGAA   4448

(2) INFORMATION FOR SEQ ID NO: 66:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 3219 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: double
       (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

CGCGCAGCGC GGTTTTCAGA TTTTGCGCCT CTTCTGCACT GAATCCACTG ATACTCCCTG   60

AGCCCGCTGT TATCGGCTCC CGGATAGCmG GGGCAGACCT aATTTTACCG TCAGAGACAA   120

TTGCAAGGCG ACGCCCTATC TCCTTAGTGG TAAGTTCCGA AAAAATACGC GCTCCTTCAT   180

GGTCCAGGTC AAACAGCACC AACGGCTCGT TCGCGCGACC TGAGCTCACC GTTGCATCAC   240

GAATATGTCT TCCTTCAAGC GCAGGCTCCT TCTTAACAAC CAGAAACCCG TCGCGCACAT   300

CAAGTCCGTA GgAATCCTTG CGATATACTC CGAGCACACT GGTGTGCTCA GGAACCAGAG   360

ACAGATCGTG CAACTGATGC GCAgskTCGA AGGTaCccTG CGGGTTATTG CGATAGTGAT   420

CGAGAAGCTT TTGAGTCGCA TCATCATCCA CGAGATGAAA CGCCAGGACA CCACGACCCA   480

TGACGATAGA ATGAACACGG TCACGGTCAG TAAGACCAGG AATCTCCACA TACACGCGAT   540

CTTCCCCTTG cCTCCGAATA ACGGGCTCAG AAAGACCAAA GCGATTAATA CGaTTCTCAA   600

GGGTACTAAG CACCAGCGCC ATCGCTTCGC TGCGTATTGC GGCGCGCTCT GCATCCGGAA   660

CTCCCTTGGT AACTTCGCTC AAATCAGCTT TAATCACCAC GCTAGnGCCG CCGGAAAGAT   720

CAAGCCCGAG CTTGACAGCT TGGGCCTGTC TCCTTTTCAT CTTCAGGACG GCTTCCCGGT   780

ACGTCTGcTC CATCAACGGT CGCGCGTAAA GAACAAAACC CTGCTCGCTT TTTACGGGGA   840

ATGCAGAAAC GAGTGCCGCA GCGGTCCAGC GAGAAGGGGC CGGTCTGCCC GAATAAGAAA   900

GATTCTGGCG CGCAGGcAaC AAGCGGTGCA TAACGCGCTG AGATATCCTC ATCCGACCCC   960

GCACGCGCAA GACGGGTTAA ATCCGCAAGA TCACGCTCAG CACTCTGCAC AGCGTACTCT   1020

TTTATCTGCT CGCGCGAGCT GAGCGCACGC TGCCGCGTTT GTGCGTCGGT CAGAAAATAC   1080

CACTGGAGTG TAGGGAACAA AAACCCAGAG CACGCAGCAA GAACAACAAG CACGACCCCA   1140

AACCGAGCCT TCTTACTCAC CTGGCGATCT CCTTGTCCAC ACCTGTCAGG GGCACGCCGG   1200

GCTTCGAATC GCAATCTGTC TTAGGATTCG AAACACCTCT CCTGTCGTTT ATGCGCGCAA   1260

TCGCACTGCG GCTGACTTCG AGCGTGCCAT GCTCATTCAC CTTTATGACA AGGCTGTGCT   1320

CCCGCACCAC GCTTACCACC CCGTGGATAC CGCCGATAGT AACGACAGGA TCACCCTTTT   1380

TTATGTTCTT AATAAGAGCC TGCGTCCTTT TCTGTTCCCG CAGATTAGGC GCAAAAACAA   1440

AAAGGTAAAA GATCAGACAT ACGACGCCGA TAGCGAGCGG TGGGATCCAG CCACCGTTCG   1500

CCGTAGTGAT TTGCAAAAGA GTTCGATGGG GCATTGTTTT CATCCTTGAG CGCGCAGGAC   1560

ACACGAGCGC GCCCCCAGGC TAGCGCAAAA AAGACAATCC AGTCAATCAC ATCTCTTCTT   1620

TACCAaCGCG CGyGyGCGCT gGCATTAATC TCAAAACGAA TCCATATCGG GCAACTCTAA   1680

AATCAGCTGG ACTTCACCCT TAGAAATCTT CAGTGCGTGC GCAATAGCGT CGTCAGACCA   1740

GCCACTTTTG TGCAGCTTCA CCACATTCTG ACGTGTAGCC AGGGGCGGCG CTCCCGCACC   1800

GGGTATTTTG TTTGCTGGAT CCTGACGCAT CAAATCACCC AGCAAGCGCA ACTGCCCCTC   1860

AGACACCTTA GAAATTTCTT GCAGACGAGT TTCAGTACCC GCAAGCCACT CACGCGCATG   1920

CTGTATCTTT TCAATACGAC TTTCCATTTC TCCCAGCAGC GCATCTGCGC ACTCTATGCG   1980

CGAGCGCACA CGCTCTGCCT TTTCCTGGTT ATCCAACAAT ACGGCAATTT CTGCACGCAC   2040

ACGCTGCAAT TGAGGATCTA CTGTCTCCAA TTCTCCCCTA AAATTTTTAA GCGTCTTTTC   2100

AAGTTCCTTC AAGTTTTCAA ACGCTCTATC CACGTCCTGC ACCGTCTGAT CCAACACCGC   2160

ACCTTTCTTA TCCAAACGCT CATAACGCGT ACTGATATCG CCAAGCCCTT CTTTGACCTT   2220

TCGAATTTGC ACCTGATAGC GCTGCAGATC GTCATTCGCT AACGTGAGCT CAACAATCTT   2280

CTTGTCCATA GCATCAGAAA GAGCAGCAAG CTTTGAAAAC TCTCCCTCCA AAAGATCAAT   2340

ATTCTTTCGC TCCTGCATAA ACTTCTCAAC ACGCTGctCC GCCTCCTCTC CAAAGTGTTT   2400

CACTTTTTCA TACTGGAGAC TAAGCTTATC CATCGCCTCT CGATACACTT CGAAACGGGT   2460

CACCGTCTCA GTCAGCCGCT CGATATCCTT TTCAAGATTC TCCCGCAACT CGTCCGCCCG   2520

ATCAAAAATA CGAGTCTGGC CGATAAACTC ATGCTGTTTA CGTTCAATCT CCTGCAGCAC   2580

TTGCGAAAAG CGGTCACTTT CTCCCTGTAA TTTAGTGAGA AGACCTACCT GCGCCTCCCC   2640

AAACTCCGCG CGCAAATCCT GCACAAGGTC TCGTGTCTCC TGCAAGGTGC GGTCCACCTG   2700

AGCACGTGCT TCTTTTACTG TCTTGTGCAG TGCCTCACTT TCGCGCCGAC CATCTGCATG   2760

GAGCTGTCCT GTCACAGTAT TTACATAGCC ACCCAGTTCC TCTTTGACCG TCCGCATCGT   2820

GTCACACATC TTGCCTATTT CATCACGCAG ACTTTGCAAA GACTCACCAT TCTTCTGAGA   2880

AAAATCTTCA TACTGCATAT CATAGCGCGC ACTCAGGTTT TCAATCGCCC GCTCAGAAAG   2940

ATTCACCAAA TGCGCAATTT TTCCTTCAAA CAACTGCTTT GCATCCgCAA ACTGCTTGTC   3000

GGTGTGTGCC TTCCATGCCT CGATATCCCG TTTGACCGAC CCACAGCCCC CCTGCGCCTC   3060

CTGCTTTATC TCCTGCACGA GCACATTCAT ATCCCGaACT TCGGTTTCAA TCATACTTGA   3120

GTGAACATGC AAAGAGTCAC GCAGgCGTTG CCgCACTGCT TCCAATTCTC TCTCAATGAG   3180

ATTATGCGCC TTATGTGCAA GGCCGCGACA TCTAAGGTA   3219

(2) INFORMATION FOR SEQ ID NO: 67:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2725 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

CAGGATnCCC CATTCCTGAG AAGAAGGCgC GCATCrCGmA GtCgACTGAC TACCCTTCCG   60

GCAGCCTCCG GTGCATCGTG CCTCACCTTT TTTACCCGTG GACACATACC CCAATTGCGC   120

AtTTCAAAAA GTCCGTTGAA CAATCGTTCG TCGTTTTCTT ACACGCAGAT GTGCAACAAC   180

TACGAACGCA AAACATCACG TGGCTTGGAT CCATTCGGCG GACCGACCAC CCCCCTTGCT   240

TCCATTTCTT CGATTAGGCG CGCGGCGcgA TTGTAGCCTA TCTTCAATTT ACGTTGCACA   300

TACGATGTGG ACGCTTTACC CGCGTATTGC ACTACCTGCA CTGCCTGCTC GTATAAAGGA   360

TCGCTTTCAT CCACAAAATT TCCAGATATA CTCGCGTCGT CATCGTCAAA GAAAATTTCT   420

TCATCAAGAT ACTCAGGCGT TCCCCACGCG CGTACATGGG CGATCACGCG CGCTAATTCT   480

CGCTCGGAAA CATACGCACC TTGAATCCGC GTAGGAAAAG ACTGACTCGG GTTCATGTAC   540

AGCATATCCC CTCGTCCCAG CAATTTTTCT GCGCCCATCT CATCCAAAAT AATACGGCTA   600

TCCATTTTAG ATGAAACCAT AAAGGCAATT CTGCTTGGAA TATTTGCCTT AATAAGGCCG   660

GTGATGACAT CGATTGACGG TCGCTGGGTG GCAAGTACCA AATGGATGCC TACTGCACGG   720

CTCATCGCGC ACAAACGCGC AACACTCGTT TCTAATTCTT TGCCAGAGGC AACCATTAAG   780

TCTGCAAATT CATCAATGAT AATAACGATG AATGGGAGAG CTGCGTGGC GATGCTTTTT   840

```
TCCTGTATTT TTTTGTTGTA GGTCTTAATG TCGCGGCATT CTAATTGCTC AAGAAGCGCA    900

TAGCGTcGCT CCATTTCGCA CAGGATGTAC TGTAGTGCTT GGAGTGCTCT TTTGGGCTCA    960

GTGATGACAG GAGTGAGAAG GTGGGCGATA TCGTTGTAGA GCTTTAACTC TACGATTTTT   1020

GGATCAATGA GCAGAAGTTT GGTTTCGTCA GGACACTTGT GGTACAGGAT AGAGAGAATG   1080

AGCGCGTTTA CGCATACTGA TTTACCCGAC CCAGTTGCGC CTGCAATGAG CAGGTGAGGT   1140

GTTTGGGCAA GGTCGATAAC CTGTGGTTCG CCGGTAACGT CTTTGCCAAG GATGACAGGG   1200

ATGGCCATAC GGTTGCTGCC AGCTGTGCGC GTATGGAGCA GTTCTTTGAA TGTAACGAGG   1260

GATCGTTTTT TGTTAGGGAC TtTCCmCCCT ATGGCGTGTT TTCCAGGAAT GGGAGCGACG   1320

ATGCGCACGC TTGAAGCAGC AAGCTTGAGC GCAACGTTGT CCTGCAGATT TGTAATTTTT   1380

GACAGTTTGA TGCCGGGTGG AGGGAGAAGC TCGAACATTG TGACTACAGG ACCCTTCTTG   1440

ATACCGGTGA TTTCTACTCG AATGTTGAAT TCAGAGAATG TTTCCTCAAG CAGGAGTGCA   1500

AGATTCTTGG TGAGCTCGTC AATTCCTTCA TATGTGTCCT CTGAGTACTG GTCAAGCAAG   1560

TCGTACGGTA CTTGGTAGCC GCGGCAAGGG TgCCGAAgCG GAGCTGCTGA GGCAGGAATA   1620

GGACGCGgTG GTCCCTGTTC ATCGTCTTGC GCAGgAATAA GGGTTTCAGC GGGGGCGACT   1680

GAGGGAGCAG AGATAGGAGA GAGGGCCATG ACACACGGTG CCTGTGCAGG GATGACAGAC   1740

GGCAAAGACC CGGGTGACGC GGGGGCGTGA ACGTCTGAGG GAAGGTTACT CTGAATGAGC   1800

CCGGGCGCTG GGAGCAGGGG GAATGGAGCC TGAGAAGGCA CCGATGGCGC CAAAGCAGTT   1860

GGCGTGGACA CACCGCCACA CGCTGCCACt GCGTGGCAGG CTGCACCTCT GCCTCAGAAA   1920

TCAAAAATTC TCCCCCCTGG AGGGGAACTT CCGTGGAGAA TTGCCCCTCT GGGGGCGCGC   1980

TTGCTTCGGG CGTCTGCACA TCTGCGGTGG CGCAGGAGGG AGCGGGGGA GGGGAAACGG   2040

TGTCAGGATG ATCGGCGGTG GAGGGAGGGA AGGAGGGGTC TTGGAATCCA TCAGCGATGA   2100

AATCcGAGGG ATACGTGCAT GAAACCATAC GTAACACcTT TCCCCGTAAA TGAGTGCAGC   2160

ATAGAGCTCT GCTCCCAGCA ATGCGAGGAG GCAAAGGACG CATACGATGT CTATCCCTCC   2220

CCGCGTTGAC GGTGAAATTG ACCGTGCcAG CGAGTGCGCG TCGTAGCGCG TAGAGACCGT   2280

GTTCTCCACA CACTGCAGTA ATGAACAAGA GTGGGAAGGC AACAAGTGCG CTTTCTGCCC   2340

GTAACGAACG TCCGCCGACA AACAAGAGGA GCGCTGTGTG CAAGAGTAAC AGCGGCACGA   2400

GCAAGGAGGA GAAAGCGTAC GTTTCGTAAA GGAGAGTGCC AGGTACGAAG AACCAGTGTG   2460

ATGCTCGGTG CAAGGTAAAA AGGGGAAGAA ACGTGGACAG GGTCaGGAGC ACTGCACTGA   2520

CGAATAGCAG TGTGCCGAAG gTAAGAGCGA TAATTCTAGG TAAAGGGGAT CGTTCCATGC   2580
```

ATTGTCCTGA ACAGTTAATC TGTTAGCTTG CACGCCCTGC AGGCTACCGA CCCCGACAGA   2640

AGGAGCCGAG TGAGGGGAGg AAACAGGCGC GACCCAAtAT CTTTGTAACG GTAAGATGCT   2700

TTGCGTTACA CTGnGACGGG CGTnG   2725

(2) INFORMATION FOR SEQ ID NO: 68:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3406 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

CGGCGCATAC TGTACCGCAT CCTCCTGGTA TGCGCTnTCA CCCACGGnTT CCACAGCCGG   60

GCAAAGTTCG TCAGAAACGT GGGGTTCTTC CCAATCGTCA GGTAGGCCCC ATAACAGTGT   120

AGTGTCGCCT CTACnTTCCC CTTGCGCTTA ACGGCAAAAn CTGCnTTCCC CTGACTCAGG   180

TCCGCCTGCA GGTCCGCCAC CTTnCAGTCC GCATACAGTG CCGGGTGCTG CCCACGGCGC   240

GTGTGGGTGG TGCGCATAAn CAGGGGAAAG GATACTCCCA CCGTGTTGGT AGTACGAAAC   300

CCGTGCTTCA GATTGTAGGG ACCGGTGCCC ATAACTGCAC CAGGGGCCTG GCCATGACTG   360

CCTACCCCCT TGCCATAGCT GATGCCCCAC TCAAGTGTGG CAGAGCCAGT TAGCTTCGGG   420

GAAAACTCCT GTCCGAGCAC TCCCCCGCTC GCTCCTACCC CCACCACCAC ACACAGCACA   480

CTCCCCCACC GCATGCACCC CATGCTACnT CACCCCCCCC CnGGnCCTGT CTAGTAGCCC   540

CyTCAcCTTC TTTTCTAACA CTACTGCCCA ATCAAGGTAT CCAGCTGCTT AGACAGCGCA   600

CGGTgTGCGC ATTGTTCGGA TCCAGTGCAa TCACCTGATG CAAATAGTAC TGCGCTTTGC   660

GGAAATCCTT TTTCTTTCGG TACCATTCAT ATAACGCAAA AAGAGTGCGT CCATTGCGCG   720

GATCTGAAAG CAAACTTGCA CGCAATAAAC TTAATCGCTC CTCCTCGTGC ACTGACAACA   780

ACGCTTCATA GTACGAAAAT ATCGACCGCA GCGTACCACG CGCGGACGCA CTCCGAGCTG   840

CAATCACCGC ACGGATCTCC CGGTAATGAC GCGCCTCATA CAACGTATCT AAGTACAAGA   900

CGATGACCGT CTCAGACGGA GGCTGTGCCG AATGATATAA ACGCCGCGCA AGAGAAATCG   960

CCTCcTGCGC ACGACCTGAA CCACTGTACG CGCGAATAAG TAATTCTTGA TGAGCCTCAC   1020

TCGGATATGC GGTATTCAAA CgCTCCGCGC GCGACACTGC CTGTTCCCAG TTACCCTGCG   1080

CCAGTTCATA CTGCGTCAAT AAACGGAGCG CTTGAGCGTT ATGCGCATCG GCGCGGAgCA   1140

CCAGCGCGAT AAAGCTTcGC ATGTTTTTTT GTGCAAGCGA GTGACCGGTC TCAAAGCAGT   1200

```
GCTGCGCACA CGCAAGGAGC ACCTGCACAT CACGCGGATA CAAGCGATAT GCCTGCTGCA    1260

AAAAGTCGTG CGCAGACGCG TCATTTTTGC TCCATTCCTT TGCCACCTGC GCACGCAAGA    1320

GCAAGTACGT TTTATCCGAA CTGTCACGCT CTGCAAAAGA GTCAAGGGAC GCGTGaGCCT    1380

TCTGaTACTC CCGcTGaGCC ACCAAGATAC GAATGTGTAG CAAAAGCGCC GCATTGTCCT    1440

CAGGCTGTTG ACGCAACAGC GCTGCAACGT ACGGTTGCGc GGCGCTCCAT TCTTGACGCG    1500

CGCTGTGAAT TTGCGCCTGC AAAAGCTGCA CTGCTCGGTC TTTCGGGAAA CGGTGAAGCA    1560

ATAGGCGCGC AATGGGCAGT GCCAGGGCAG TCTCACCGCG CTGTACCAGC ATGCGCGCAT    1620

AGCAGATGCC TGCAGGATAG CAGGATGCGT CTTTTTCCCA TGCGCTGCGG TACCTGnACT    1680

CCGCTTCAGA GAGCTCTCCG TGCTGCTCGT GCAAATACCC AAAGAGCAGG AACGGAAGGA    1740

GAGAGTGCGG GGCCTGCGTG CGTGCGCGCG TtCAGGCGTT CTCGGCAATC TGCGTGTACC    1800

TCATCAGGAA GGCGTTTTTT CAACAGAGTC AGCGGGGGAA TGATATCCGA AAAAAAATCC    1860

GGTTCTCCAG GAATGAGCGG ATACGTACCT TTTTCCACGT CATCGAGCGC AGTCAGGTAC    1920

GGATGCGCGG TGTTGTACAG CGGCACATGC CAAGAAACAG CTTCCTGCGG ATATACAAGT    1980

TGCATAAGCG CAnACACACG CGGAGGTACA GCCGATTGTG CGGTGTCAAT CCGGCCGGAT    2040

CGCGCTGTAT GCACGCAGCT GCCTCGCGCA ATGAGGCAGG AGAAGCAGTT TCAATCAAGA    2100

AGAGTATCTT TGGATCCATC AGCTTTGTGC GAATTGCGCG TCGGTCCGGT ACGTCAAGCG    2160

CAtGCCaCGC ACAGGaTGCG CCGCCACCTC CGGCGCCTGG GATGCGGAAG AAACCGGATC    2220

GGTACtGCGC GCGTCGGCAC CACCTTCGAA GAACTGCGAC AACAGAGAAA AACGGGTATC    2280

AGCACAAAAC CGACACTCAC CCCGATGGCG CGGTCAATGC TTTTAGTAAG CGCCCCCATA    2340

GAAACCCGTT TTTACCTCCc TGcATGGCAG TCGTGCCATT TACACAAAAC GCCTGTTGTG    2400

ACCGTACCGA CAACGTACGC ATACCGGCGC ACCGCCGCTT CTTTTCCCTT TACGCGGTGT    2460

TCAACGcgCA CGGCGCATCC CTTGTGCTCC CCACGCAAGA CATGCTAGGC TGGCTGCCAC    2520

CGAGGGCGAA GAGAGCGTAC AGGAGGTTAA CGGTTTTTTT GCGAGAAATC ATTACCGCCC    2580

GTGCGTGTTC ACTCTTCCTG TTTCTCCTTC GTTGTTTCCC TGCTGGTCCC TGTGCGCGGG    2640

CGCGCCGGTT GTCCTTCCTT CTGTTTCTCT GTGGTGCGGC AGCCTGCCCT CCGCTTTGGG    2700

GGGCGTACGC AGCGCACCAg CGTTGCGCGC TCAGTCGGTA CCTGACACCC TCATTCAGCG    2760

CGCGCTCGTG CTCGGTCCGC TCGTGCACCC CCTGTACCCG CCGATGCAGT CCTTCAAAGA    2820

ACAGTACCGG AGCGCGCGTT ACCGGGAATA CCTCTCTGTC GTTATGcAGC GGAGCGCGCC    2880

CTACCGCCCC TTTATCGAAA AACTGTGCGC GACGCTCACC TTCCTGTCGA .GCTGCTCTTT   2940
```

CTCCCCGTTG TCGAATCGGG CTTTCTCGAA CGGGCTGTCT CCAAATCCGG CGCAGTCGGC   3000

ATTTGGCAGT TCATGCGCAA TAGCATCGCA GGATCTGCCA TGCGCGTGAG TGACTGGGTA   3060

GACGAACGGC GTGACCCCTG GAAGGCTTCC GTCGCCGCAG TCAAAAAACT GCAGTGGAAT   3120

TACACGCAGC TGCGTGACTG GCCCTTGGCC CTCGCTGCGT ACAACTGCGG TCTTGGCGCG   3180

ATCAAGCGAG CCATTGCCCA GGCAGGAACC GCCGATTTTT GGCATCTGAG TGAGCGCGGc   3240

TTTCTGCGCG ACGAGACAGT CCGCTATGTC CCAAAGTTCC TTGCGGTTGC AGAAGTACTC   3300

AGCCGGAGCC ACGAGCACGG CATCGCCTGG GGAGCGGCAC ACACCCCCGA GGAGACCACC   3360

ACGGTTACCG TTTCGCGCGC GGTAGACTTA AACCTCTTGG CACAGG   3406

(2) INFORMATION FOR SEQ ID NO: 69:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 7874 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: double
          (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

TGATAGCAAA TTATCTGCTg AAAGGCTCAC AGTACAGCAT GTTGGTCCCA GTGGTTGTTC   60

GCCTGGGGGC GtATACAAGG TGTGAGGGAG TTGGCATTCG GGGGGCGTGT GCGGAAATGA   120

AATGGaGTGG GCTCTGTCTC TTTCTCCGGA CGGGGGGGGG GGGGTGCGAA CAGATGAACG   180

GAAAGCGTGT GTTTCTGGGC ATTGTCgTGG TGGTCtGTGC CGCGCGCTGT TTTTGCGCcG   240

GACGTGTTCT TCTCCTCGCa TCTGGGGTaT GGGCGTTTCT ACGCCcGTGG GGAAAGACAT   300

TGAaGGGGCa GCACATGCAC GTTCCAAGCA TTGGCGGCGG CGTATGTGTG GTGGCArACA   360

GCGGGTTTGC CTTCGCCTGC ACGGTGGACG CAGCCCTGAC CCGTATAATG CTGAAAACTC   420

AGGCGCTCTT TGGCTATGCC TTTCGGTGGG GAGCGTTCAG CCTCATCCCC TTGCTTGGGA   480

TGGATGTGAT TGTGTCGAGC GACCACGCGT TTGGTGTTGC CGCGCAAGTG TCGTTCCAGC   540

ATTTGATTTC TGAGTGGTGG GGCTTTGCCT TGAGTGTGAG CGGCGGGGTG GACTTTCCGC   600

TCAACCCTAA CACCCGCTTT TTAGCAGGTA AGCTGCCTGC AGAAACGGTG CAGCGCGTGG   660

CsTCGTTGCG CTGCGGCAAA AGCTTATTAG CGAAnGGATT ATCAAGGCAT TGGATTTGGG   720

CTGGTTTATT ACCTTCGCTC TGACCGTTGT TGCCGAGGGA TTCAGTTGGA TTGTGTCGCA   780

GAGCGCTTGG ATTGCGCAGA AGGCGGTGAA TTACTTTTTG AGCGACACCA CGCGTTGTCT   840

CATTCTCCCG GTCACGCTGC GGGCCGGTCC TACCTTTCGA ATATAGCGTG CGGGGGGGGG   900

●

```
CGGATTTAGC GGCGCGTTGG CGTCCGCTGT CGAGTTGCCC AGAGCGCGAG AAGAATGCGG    960

TCGATTTCTT GTGTAGCGTT CCGTCGTGCG CCGCGCACTT GAAcGAGCTC AGCCGGTGCA   1020

TGAGCGAAAs CgCGTGGACG ACTGTCTGCC CGAACTTTCC CGGCAGtGCG GCGGCACcGC   1080

AGTATCCACT TTGAAAGGTA TACCATCCCC GAGGCCATGC AGGCCTGGCG CGCTCCTGCC   1140

GAAAGCGCAC CGCTCGTCAG GGGGACCCCA AGCAmCGTTG CCCACACGCG TCCAACCGGC   1200

GGTCTAAGGT AGCAGGCGAG CGTTCCCACC CACAGCCCTG CCAAGGTCAT CCAGCGGATG   1260

CGTCCCCCCk TTGCCGCATA GAGCAGCAtG CGATCCCGCA CACGGACGCC TGCACGCTCA   1320

GCCGCGGCAC ACACAGCAGT ATCAACAGGA GCAGGAGTCC TGTCCCCATC CGTACATAAC   1380

GCAGCACACG CGGCCTAAGC CgTGTTCCGA ACACGGCGCC GTGCATTCAC ACCCCGCCTG   1440

TACCGGAAAA AAAAGAGACG CAGGCGGTGC GGTGTCTTGT TGTCCGCATG CCGCGTTCTG   1500

CAGGTGCGCG TACCACTGGG AGTGGGAGAG AAAATGGCCG GTTAGCGCTC CGAGTAAGGT   1560

ACTGCTCACT GCCCCCAGCG CGCAAAATAG CGGGAGTACG TACCACACCC CCGCGCCAAA   1620

AACGAGCACC CGAGCAAGCG TAAGCTGTAT CACGTTTGAA CAGAACGCAC CCATCACGCT   1680

TATCCCCACG CACGAAAGGT ACCGGCACGG GACGAACCGC AGCGCATACA TGAGCGCACC   1740

CGAAGCGGTG CTCCCTGCAA GCGAGAGGAC AAATACATAA GAAAAGAGCG TCCCACTCAC   1800

CAGAGCCTGC CCTATTACCT TCAGGAATAC TAAACGCGCG TACGCACAGA AAGGGAGCAG   1860

ATCCGGCGAG ATCAACAGCG GCAAATTCGC AAGCCCCACG CGAAAGAAAG GCAGCGGCTT   1920

TGGAATGACG TGTTCAACCG TAGAGAGAAA GAAACACATG CCGCCTAAAA GCGACACTAA   1980

CTCATCGCGT ACGTCTAGTG GCAGCCTGCT CCGCACGAGC CGCACCCTCC CGCGCCGCTC   2040

CCACAGCCAC CGCCGCTGGT GCTTTCCCGA AACAGAAGTG CAGCTAAGTC ATCGTCCGTC   2100

GCCTCTCGCA CAGAGCGCAC AGCCACCTCA AAGTGGAGCG TCTTCCCCGC GAGGGgATGA   2160

TTTCCATCTA CAATAATCGT TTCACCTTGC ACGTCAGTGA CGGTCACCGG TCGACTGTCA   2220

CCCCCGCTTC CTGCATCAAA CCGCATGCCC ACCTCTATTG GCACGTTTGG AGGAAACTGA   2280

TCTCGCCCCA CTGTCATGCG CAAGTCCTCC TGCACCTCTC CATACGCTCC TACCGGAGGA   2340

ATGGTTACTG AAAACTCCTC CCCCTCTTCT CGGTTAATTA AGGCGGTCTC GAGGCCAGGA   2400

ATGATCATGC CGTGCCCCTG AACATACTCG AGCGCACCCA TCACGTCGGA AGAATCGATG   2460

ATCTCCCCCt GTCATCTCgC AGGGTGTACT CGATgTTCAC CACACACTCA TTTGCGATTT   2520

TCATGCGCGG CATGCTAGCA CAGGCAAGAT aCTCACGGCA AGGGCAGTTT CTGTGCCGTG   2580

TGCCyTTGAc AGAATCGCCG TTATAGGGGA TAAGCCGGGC GAGGTGTTGG GAGCGTGTGG   2640
```

```
TCCACTTCTT GCCCTCTTGC GCgGTGCTGT GCgGTAAAAG AgGGGGCGTC gCGTTCGAGT    2700

AAAATTTTCT CTTAAGCCTT AAGTGAGATA CCCCATTATG GTAGAGGTCt AACCGCGGTT    2760

GCGCGCTGCT GTTGCTCGGT TGGCGGTCTG TAGCGCTGCG GAGAAGGACG GTGCCCTGcC    2820

GCTGTGGCGA TGCGCTACAT GCGCAGCGGG AGGATATCCT GCGTGCAAAT GCGCAGGATC    2880

TTGCGCGGGC GCGTGAGGCG GGTCTTGCCg CACCGCTTGT CGCCCGGCTC GCGCTGAGTG    2940

AACACCTTCT TGAgGACATG TTGCGGTCTT TGaCTGTTCT TTCGCTTCAG CGGGATCCTA    3000

TCGGGGAAAT TATAGAAGGG TACACTCTTG CGAATGGACT GGAAATCCGG AAGGTACGTG    3060

TTCCTCTGGG GGTGGTGGCT GTCATCTACG AGTCTCGGCC CAACGTGACC GTAGATGCGT    3120

TTGCACTTGC GTACAAAAGC GGCAATGCGG TGCTCCTGCG CGCAGGTTCT GCAGCGAGTT    3180

ATTCAAATGC CCCGCTTTTG CGCGCAATTC ACGTGGGTTT GAAGAAAGCG CATGGTGTCG    3240

TGGACGCGGT GGCTGTTCCT CCCGTTTTGG AGGAAAAATA TGGTGATGTG GATCATATCC    3300

TCcGCGCGCG CGgCTTTATC GATGCGGTAT TTCCTCGTGG GGGGGCGGCG CTTATCCGGC    3360

GCGTCGTGGA AGGCGCCCAC GTGCCAGTTA TTGAAACCGG ATGCGGCGTG TGCCACCTAT    3420

ACGTAGATGA GAGTGCGAAT ATCGATGTGG CGCTGCAGAT TGCAGAAAAC GCGAAGTTGC    3480

AAAAACCGGC CGCATGCAAT TCAGTCGAAA CGCTGTTGGT GCATCGTGCG GTTGCGCGTC    3540

CTTTTTTGCA CCGTGTACAG GAGATTTTTG CCACCTGTGA GGAGACTACG CGCAAcCCGG    3600

TGGTGTGGAT TTTTTTTGTG ATGCTGAGTC TTTCTCCCTT CTCACAGAAA GGGGCGCGAG    3660

AAAAAATGTT TTTCATGCAC AGGCAGAGAC CTGGGATCGG GAATACCTGG ACTATCAGGT    3720

ATCCGTGCGG GTGGTGCCAA ACCTTGAAGA AGCACTCAGG CACATTGCTC GTcATTCTAC    3780

GAAACACTCA GAGGTTATTG TCACGCGCGA TCGTGCCCGT GCGCGTCGTT TTCATCAGGA    3840

AGTAGATGCT GCCTGTGTAT ATGTCAATGC TTCAAGTAGG tTTACCGATG GAGGGCAGTT    3900

TGGCATGGGA GCAGAnATTG GGGTCAGTAC GCAAAAATTG CACGCGCGCG GTCCGATGGG    3960

TTTGTGTGCA CTGACTACTT CAAAATATCT GATTGATGGA GAGGGGCAGG TGCGTCCGTG    4020

ATCCGTGCGC TTTTTGCTGC GGCAAAAAAA AtTGTGATAA AGATTGGGTC AAATACGCTT    4080

GCGCAkGCAG ATGGTACTCC TGATGAGGAG TTTTTGGCGG wGTGTGCTCG CGCCTGTGCG    4140

GCGCTGATGC GTGACGGCAA GCAGATAGTT GTGGTGTCGT CTGGCGCTCA GGTTGCAGGG    4200

ATTTCTGCGC TCCATTGCCT TTCATCTCCT CCTCAGGGGG CGGGTTTAGA GCGTCACGAA    4260

TCGCGCGGCG TTATTCCGGG TGATGGTGCG TCCTGCAAAC AGGCGTTGTG TGCGGTGGGT    4320

CAGGCGGAgT TGATAAGTCG TtGGCGTTCT GCGTTTGCAG CGCACCAGCA GTGCgTGGGC    4380
```

```
CAGTTTCTGT GTACGAAGGA GGATTTTACT GACTCGGACC gCGCGGCGCA GGTACGCTAC      4440

ACGTTGTCCT TTTTGCTCGA GCGCAGGGTA GTACCTATCC TTAATGAAAA TGACGCGCTC      4500

TGTTGCAGCG ACGTCCCCTC TGTAmCCGCC GACCGGcGGt GTCCCTATCA CCTCAAAAAA      4560

GGATTGGAGA TAATGACAGT CTGTCCGCGT TTGTAGCGCT GTTGTGGCAG GCAGATCTTT      4620

TGCTTTTGTT GAGTGACATT GACGGCGTGT ATGACAAAGA CCCAAAGGCA CACACAGATG      4680

CGCAgcACGT TCCTCTGGTG ACGGACGTGT CAGCGCTTGT GGGTAAAACG AGCATGGGTT      4740

CTTCCAATGT CTTTGGTACG GGTGGGATTG CTACAAAGCT GGATGCTGCG CGTCTTGTCA      4800

CGAGGGCGGG AATTCCTCTG GTGCTGGCAA ACGGGCGCCA TCTGGATCCG ATCCTGAGCC      4860

TTATGCGCGG GGATGCGCGG GGGACACTTT TCGTGCCTGT TTCTTAGAGA GCGACGTGGG      4920

TATGCGCAAG TGCACGCATT GTGCCCTATA ATGCGCGGCG TGCGGTCAAT TTCTGACGTG      4980

TAATTTTTCT CGGTGGGGCG ACGTCTCCGT CTGTCTGTTA ATTCGGTGGT GTGTTTCGAT      5040

GCGAGAAAAG GAAGGAGGTG TGGTGAACGA CGATTTTCAC TATGAAGTGA CGCGCAACTG      5100

GGGCACGCTT TCCACATCGG GGAATGGCTG GTCCCTCGAA CTGAAGTCTA TTTCTTGGAA      5160

TGGCCGGCCA GAGAAATATG ATATCCGCGC GTGGTCCCCA GACAAGAGCA AGATGGGAAA      5220

GGGGGTaACg cTTACGCGTG CAGAGATTGT AGCCCTGCGC GATTTACTAA ACAGTATGTC      5280

CCTGGACCCG TACTAGGGAC AGTCTGCAGT GCTTTGTGCA GcGCGGCGCg cAGcgTCGGt      5340

GGCTAGCCGG TCGCACAGTT CGTTGTACGG GTCTCCTGCA TGTCCTTTTA CCCAGCGCCA      5400

CTCGACGGAT AGGGCGTCGG CGAGTGCGCT GAGCGCTTCC CACAAATCCT TGTTCTTGAC      5460

CGGTTGTTTG GCAGCCGTTT TCCAGCCGTT GTGTTTCCAG GTATGGATCC ACTGGGTGAT      5520

GCCTTTGCGT ACGTATTGGG AGTCGGTGAC CACTACCACC GCCTCTGCAG CGCGTCCGTG      5580

TGCCTCTTGC AGTGCGTTGA TGACCGCGCA CAGTTCCATG CGATTGTTtg TGCTCGGGTA      5640

GGCgCTGCCG CTTCTAGTGA ATGCGGCAGC TTCTGGTGCG GTTTGTCCGG TTTCTAGAAA      5700

GGGTACGTCT GAGGGCACCA GAGCAAACGC CCACCCGCCC GGACCCGGGT TTCCCAGACA      5760

GGCGCCGTCA GTGTACAGGG TAAGTGCAGC GTGCGCGTTC ATAGTCGCGC tACGGTAACA      5820

GTTTTGCGCC GTGGGGACAA TGTATTGGTC CGACAGTTGG TGATGGAGCG AAGATATTTT      5880

CGCAAGGAGG GAGAATGAGG CGCGCACGGA TTGTGCAGGA ACTTTGGTAC GCGGGACGAC      5940

GGTTTGGTTT TTGCGGTACG CTGTCTTATT CTGCAAGGCG GTGTACACGT GCGCGTTGCA      6000

CTTTCTCCTC GGGTGTACAT GCTGCACTGT TTTTAGAGGA AAGCTAACAC GGAGAGGGCA      6060

CAGATGAATA TTCTGCATAA CTTTGTTGTA TTCGAAGGTA TTGATGGCAC AGGCACGAGT      6120
```

```
ACACAGTTGC GTGCGCTCGA ACGCCATTTT CAGGCCCGTA AGGACATGGT CTTTACTCAA     6180

GAGCCTACCG GAGGGGAGAT TGGCACTCTC ATTCGGGATG TGCTGCAAAA GCGTGTGATC     6240

ATGAGCTCTA AGGCATTGGG ATTGCTCTTT GCCGCAGATA GACACGAGCA CTTGGAAGGT     6300

GCAGGAGGCA TTAACGATTG TCTTGCAGAA GGAAAGATAG TGCTCTGCGA TCGGTATGTT     6360

TTTTCCAGTT TGGTGTACCA AGGCATGGCG GTGTCGGGTA GTTTCGCGTA TGAATTAAAT     6420

AAAGAGTTTC CGCTTCCTGA AGTTGTGTTC TATTTTGACG CGCCTATCGA AGTATGTGTT     6480

GAGCGTATCA CCGCACGTGG GCTGCAAACG GAACTGTATG AGTACACGTC TTTTCAAGAA     6540

AAGGCGCGCA AGGGGTATGA AACTATATTT CGCaAGTGCC gTCaTTTGTA CCCTGCAATG     6600

AAAGTGATTG AAATAGACGC GCGCGAGGAA ATTGAAgTTG TGCATGAGCg TATTCTTCAC     6660

CATCTGCGCG AATACAGGCG TCTAAAATAG TGTGTGGACG TAGATACACT ATCTGAGGAG     6720

CAGTGGAGAG TATATATCAG GAACGTGCTT TGCAAGCGGA AGGCGCGTGC TCGGTAAAAC     6780

GGTGCTGCAC CGGCGCAgcA TaAGCAAAAT AATTGGAAAA TTTGTCCATA GGTTTTTGTC     6840

GTCCGGTCAC AGTGCTCAGT GCCTTTTTCT AGGCTGTTTT TCAATAACTG TTTATGTAGA     6900

CTGGACGGGT CTTCCTTTCT CAACTCACAT ATTCTTTTCG GGGACATGCT GCCGTTGGCA     6960

GACGTTGGGT GTGACGGGTG TTTCTCTGGT GTGTAAGAGG AAGATATATT CCCCTTTTGT     7020

ATCTGCACTG ACCCCTGCAC GGGGTACAGG CTATTGACGC TTCCTTTCGT CTGTGTGTCT     7080

TCACTGTTGC GTGTACGGCG CGTGAACGGG CCATATAGAT AGATGCTTGA CGGGGTCTGG     7140

TTGCCATGTT AGGATCCACC AAGCGTGACT ATTCTTTTCT GGCCGCGTGT GATGCATAAG     7200

ACACTCCCAT AGCACCGTTA AGAGTCTCGC GAAACCTCCT CCGTATGGAG AGGGGTAATC     7260

CAATTGCCGT GGAACGCGAA GGTTCTGTGT TATGTCCGCA AAGATTTACG TCGGTAATTT     7320

AAATTATGCC ACCACTGAGG CTGGATTGGC CTCCCTTTTT TCTCAGTTTG GGGAAGTGCT     7380

GTCCGTGGCT GTAATCAAGG ATAAGCTTAC GCAGCGGTCG AAGGGCTTTG GTTTTGTTGA     7440

GATGGAAAGC GCAGAATCAG CCGAGTTGGT TaTTAACGAG TTGAATGAGA AGGAGTTTGA     7500

AGGGCGTAnG CTTCGCGTTa ACTATGCGGA GGAGAAGCCG CGTTTTcCCT TTaAGAATTA     7560

GTGGAGGATG GGGAGGACTT TcCATCGTGG CGCATGTTTT TgGCGTAAGG TGCTTTCGCG     7620

TGCGTTaTCT CATTTcTCGT CGTCTTTTGG TTcTCCCCGT TTGTGTGCGT CGCGGTgTGT     7680

TTGGTTcCTG TTaGGAACCC CTTCGGGGcT TCTGTcTATT TTGcTCCCAA GACTGCTAkT     7740

ACTATGGaTG agGcTGcGTC TCGCGyCCCA GGGTTgyCaw GwAgGGTGCC gTCtTTTGCG     7800

CCTGGGTTGA AGCAAGGTTT GCCnGGAACG TTGGGTCCGT TGGGTTGAAC CCAAGAAAGA     7860
```

AAAAAGTTnG GGCC

7874

(2) INFORMATION FOR SEQ ID NO: 70:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20682 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

```
GTTATGGTCC CTGTTATTGG CGATTTAAGG ATGCTGCcgT ACGTGCrCTA TCTTAtmCGA      60

AasTGCgCTG CGATGCgGTT TTTcATGCTG CCGCGTATAA GCaCGTTCCT ATGATGGAAC     120

TCAATCCTGT TTCAGTGATT GAAAATAATG TCTTCGGCAC CAAATTCTTG CTCGATGCCT     180

GTATTGCGTG TAGGGTTAAG CGCTTTGTAC TTTTGTCCAC TGACAAGGCC GTGGATCCTG     240

TTTCTATCTA CGGAGTATCT AAGATGCTCA ACGAGAAGAA TGTCTTGTAT GCTGCTGAGC     300

GTGTGCGCGA TTTCGGTCAC GATGCCGCGT ATATGTTTGT CCGTTTTGGA AACGTATTGG     360

GTTCCCGTGG TTCTATCATG CCGCTCTTTA TTGAACAAAT AAAGAAAGGG GGGCCCGTTA     420

CCGTGACAGA TCCTGCCATG ACACGATTCT TTATGACTAT TCCCGAAGCG TGTTCACTCG     480

TTTTGCAAGT CGGTGGAGTA GGAGTAAATG GAGCGTCGTA TCTTTTGGAC ATGGGGGAGC     540

CTGTGAGCAT TATGGAGACT GCGCAgcAAC TTATTCGCTA TTTTGGTTAC GAGCCAGACA     600

GAGATATTCC TATCCACGTG GTGGGCTTGC GTCCTGGCGA GCGTCTCAGT GAGCCACTCG     660

TTTCCAAAGA CGAGCGTATA GAGCCGACGG TATATCCAAA GGTTCTGCGT TTGCGTGAAC     720

GTGAACCTTT GGATTTTGCG CACCTTGAAC GCCTGTGGGA TCAACTGTAT CCTTACTGTT     780

TCCCTTCAGG AGAAAAGGTG CGGTACCGGC ACAAAGAAGG ACTTGTCCGC GTGCTATGCG     840

ACTCGTGCGC GACACTGAAA CAGCGGTATA TGCCAAATAG CGAGGCATAG GAAAATGGAA     900

GGTACCGTGA AAAAAAAGAA AGAGGGTGTT CGTGATGATA ACGCGCAGCA TGCGGTGTTC     960

AACAAACAAG TGCCGTTTTT TGTGCCCTCG TTTTCTGAAG CGGAAGAGCG CGCAGTCTGC    1020

GATGTGTTGC GTTCAGGATG GATTACGACG GGAACACAAG CACTCGCGTT TGAAAAAGAG    1080

TTTGCckTwT gTGGGtGCTC CCTATGCGTG TGCGGTTAAC TCAGCTACCA GTGGTTTGCT    1140

TCTCACCTTT GATGCAATGG GCATTGGGCC GGATAGTAAG ATACTTACCA GTCCTTATAC    1200

GTTTGTGTCT ACGGCGAGCT CTGCACTCCA CCTAGGTGCG CAGGTGGTGT ACGCCGATAT    1260

CGAGCGCGAC TCTTATAATA TCAGTGCAGA GTGTGTTGAA GCGTGTTTAA AAAAGGATGC    1320
```

```
GCGCATCCGT GCTATTGTAC CCATCCATAT TGCCGGGAAT GTATGCAATA TGCGTGATCT      1380

CAATGCTCTT GCGCGTAAGT ATCAAGTGGC AGTGGTGGAA GATGCAGCAC ACGCTTTTCC      1440

ATCGAAGACT GCGTGTGGGT ATGCAGGCAC ACTGTCACAT GCGGGGGTAT TTTCCTTTTA      1500

TGCCACCAAG CCGTTAACCA CCGGTGAAGG AGGTATGGTT TGCACAAATG ATGCGAAgcT      1560

TGcAGCGCGT ATTGCGTGTT TGCGTTCACA TGGCATTGAC CGGGCTATTT GGGATCGGTA      1620

CACAAATGGC ACCGCACCGT GGCGTTATGA CGTAACAAGC CTTGGGTGGA AGTGTAACCT      1680

GCCGGATATT TTAGCAGCAA TTGGACGCGT ACAGTTGCAG AAGGCGGCGC ATCTTTTTGC      1740

ACAACGCGCG CGTATTGCCG CCGCGTTCAC GCGTGCTTTT TCTCGTTATG AATTTTTTTG      1800

TACTCCGCCT GATGGGGATG GAAACGCGTG GCATTTGTAT TTGTTGCGCT TAGTTCCTGG      1860

AACGCTTTCT GTTTCTCGGG ACGAGTTCGT CAGATTATTG CAGGAACGGG GATTGGGCGT      1920

TTCTATGCAT TTTATTCCTC ATTTCGAGAT GACGTTTTTT AAGAAAAGTC TGTGTGTACG      1980

AGCGGAAGAT TTCCCTGAGT GTGCGCACAA GTATCAGCAC AcGcTTACGC TTCCGTTGTG      2040

GCCGGGAATG GATGACAGTT GCGTGGCGTA TGTGATAGAG ACCGTGGTGC GCACCGCACA      2100

AGAATGTGCA AAGGGAAGAG CATATATATG AGCGTGTTCG TTTCAGACGG TGCGCGCACA      2160

GGGAGCGTCT ATGCACAGCT TGTCCGTGCG CCGCGCGTTG CAGGATTGCT GCTGAACATA      2220

GATATTCCCT CTCTCCtGAC GGGTACTCTT TTTATACTGC AGCACATATT CCCGGATGCA      2280

ATGCCGTTCG GTGTGGGGAA AATACTGTGC CGGTTTTTGC GCATGGAGAG GTGGTGTACG      2340

CAGGGaACCG GTGGGTATCC TCATTGGGCC TGATGAGCAT GTGGTACGTA ATTTAGTGCA      2400

AGATGTGGTG GTGCATACGT GCGCAGAGCG GGCCTGTGCG TCGGAAATAC TCTGTGGAAT      2460

CAGTGAAGGG GAACCCCTCG CTCAAAAGGT GGCGGTGCAA GGAGATGCAG AAACTGCTTT      2520

TAAACGCGCA TCACACACGG TATGCTCCTC TTGTACATTT GAGCCGCGTG TACACTACTT      2580

TGCGGAAATG CCAGAAGTAC AGGCACTACC CGACGCGCAC GGTCTGCACG TGTACGCTGC      2640

TACGCAtGGc CTGCGCACAT GAGAAAAACT ATCGCGCAGg TACTGAATAT TTCTGAGCAT      2700

GCGGTGCACG TACATCCGCA GCAGGAAGCG CTTTCCTGTG ATGGGAGAAT ATGGTTCCCC      2760

TCAGTGATGG CAAGTCAGGC GGCGCTTGCA GCCTATTGTG CGAAAAAGCC GGTACGCTTG      2820

TCTTTTTCCT TTCAAGAGTA TGTGCAGTAC TGTCCTAAGA CTCCCAAGAT TACCATTGCA      2880

CATCGCACGG CGCTCAACGC CGCGCATGCG GTAGAAGGTA TGTTTGTTTT TATCTCCCTC      2940

GATGCAGGAG CGGGGAATTT ATTGATCGAT CGTATGGTTG CGCATATGGT CCATACTGCA      3000

TTAGGAAATT ATGAAATTCC TCGGTACCGC ATTGAATGCA CAGCGTTTCG TTCAAATGTT      3060
```

```
GGATTAACGG ATGTTTTTAA TGGATGGGCA GATGCATACA CTTCTAATGC ATTAGAAATG    3120

CATATTAATC AGTTATGTGC TGAGCTTCAT ATATTCCCTG ACGAGTGGCG TGTGGCGCAC    3180

ATGAAAGATA CGCGGGAAAC ACAGCGTTTT GCGCGGTTGC TCGCCTATCT GTGTGAGGAA    3240

GGAGATTTTC GTCGAAAGCA CGCAGCCTTC AGCATGGTCA ATGCAGTACG AAAAGCACAT    3300

GACACCCATG CCTGGCGTGG TATTGGACTC GCGTTGGGGT TTCAATATGA TCCGTCTGCG    3360

ATGTTAGCCC GTTCGGGTTT TTCCTATGTA TTACAAATGA CGCTGCACAC TGATGCGCGC    3420

ATTGTGGTGC ACAGCGTTCC GCTTTCTGAT TCGTTTAAAC GGGTAGTGGT TGCGTTTCTC    3480

ATCAGAGAGT TTGCGTGTCT GGAGGATGCG ATCTTTTTTA AAAGTAGTGA TGAGGCGTAT    3540

GGCGTGGATC TGTTGGGTCC GTCTGTGGAA TCAGTGGGGA TGAGGGTGTT TGCACGGTTG    3600

GTGAGAAAGT GTGTACGAGC AATTCAGAGA CAACGCTTCA GAAAGCCACT TCCTATCACG    3660

GTACAGGGGT CCTTTAACAC GGCCAAGAAG GGGCAGGTGT ATCAAGTGGT GACTGTTGCT    3720

AAGTCAGATG TTTCGGTGCC CGATGCGCAA TCTGAGCAGT GTGCCTCAAA GGTACCTGTG    3780

ACTGCTGATA CTAGCGGAAA ATGTGAGGAT ATGAACGGTT TTACCAAAAT GCACGGAATG    3840

AGCACGCACA CTCCTGCAGC CTGTATTATT GAACTCGAAT TAGATGCGTT GTGCGTGCAA    3900

CCTAAGATTG TCAGGTTGTG GTTTGTTTGC GATCCTGGGT ATGTCTTTTG TGAAAAAGAT    3960

GTGTACCGTA CCGTGAGTCG AAGCATTACT CGTGCGCTTT CGCACGTATC TGTAGAAAAG    4020

ATTTGGGAGC GTGCGCGCAC ACCCGAGTAT GTTATCATCG ATCCATCCGA TACTCCTCCC    4080

TATCACGTCA CCCTTTTGAG TTCAAATGCT GCTGCGCGTG CGGTGGGAAC GGTTGCCGAA    4140

GtATTGTTCC TGCtGCGTAC TACGCAGCAC TGCGGCAAAT TTTGCCGATT TCTCAAAACG    4200

CTACCCATAA GGTTCCTTTT GTTGCGCGGG ATATTTTTTA TGAGATGTTT TCCCTCAGTG    4260

CAGACGATTC TCTATGAATA TTCGCTTTAC GTTGAATACA GAACAGGTAC ATGTGGATGC    4320

TATGCCtCAT GAGCGTCTTT CGACCGTTTT ACGGAGATGT TTTCATCTTC CTTCGATAAA    4380

AGGTTCACAC GGACATGGAG AGAATGGCGC GTCTACCATT TTGTTTAATG GGGAGGCAGT    4440

ATCCGCGTAT ATCATACCCT TTTTTCTTGC GCACGAAACA CAGATAGTTA CACTTGATTT    4500

TTTTCAAAAG ACTAAAAGAG GACGGTGGAT TGTTGCGTGC TTTGCGCAgC ATgCATATTT    4560

CTTTTTGTGG GTATTGCGAT GCAGGAAAAA TTCTCACCGC AGAGAGCTTA TTACAAAGGA    4620

ATTCTGTGCC CAACGAAGAA GAGGTAAGAC ACGCCTTTTC AGGCATGCAA TGCnGGtGTA    4680

CTGATATCAA TGCGTTGATT CGAGCACTTC AACGTATGCC TGCTTGTCAT GAGTTTTCCT    4740

AAAACTGTAT GCATCTGTAT AAGAACTCAG AATCAAGTGT TATTTATTAC GTAAAAAGTT    4800
```

```
TGGGTCAATT GTGTGCGGTA CTACGTAACG TCGCGCAGGT ACAGCCAGTT GGGGGTGGAA   4860

CGGGCTTGGT GCAATATCAG ATAACCCCGG TTTTAACGTT GCCTTCACAT TTGGTTGTTC   4920

TAAACGGTGT ACCAGAGTTG AAAGATATTT CTAAAACTGA GCACTTTCTT GAGTTTGGTG   4980

GTGCTGTGTC GTTGCAAGCA ATTGTGCGAT TGGGAAGAAA AAATATTCCC GTGGCACTGC   5040

ACGAGGCACT GTCGCACGCA GCAAATCCTG GGATACGGAC TCTGGCCACT ATTGGGGGGA   5100

ATATTGCAGG TACGCGTCCG CATGCTTCTG CTCTTGCGCC GCTTATCGCG CTTGATGCAA   5160

AAATGGAGGT GCGGACTGGA CATGAAAACT TTTGGATTTC TGTGGCACAC TATGCACATG   5220

CGCGTTCTGA CACGCTGCGA CACCGGAGTC ATGTAATTAC CCGTATTCGC CTTCCAACAG   5280

ATTACTGGGA CTTTTCCTAC TACAGACGTA TTGGGTCGCG TGCATTATTT GGTGAACGTG   5340

CCGATTTCGT GTTTCTTGCA CAGCAGCAGA AAAACGCGTT GTCTGAAATG CGTATGGTAT   5400

TTTTTTCAGA TGTAGTAATG AGAAATAGAG AATTTGACAA TTtGCTGTTA GGCAGAGCGA   5460

TTCCTCTTTC TGCAGGGGAT ATTGCGGCAA TCGTATATCG AAGCAGAGAG TTCTTTGCGC   5520

CTGAATCCTT TAAGAGTGCG TACATCGCGC ACTGCTTCTT TCATCTGCTG GAAGACTGTT   5580

TGCGCCGCTT AAGATGAAGC TACAGGTGGC GAGTTTTACC CAGGCACGCG CAAACAGcTG   5640

ACGCTAAAGA CCGAGTTTTT TCATTTTGCT TTGCAATGTA CTTGGCTTAA GACCAAGAAT   5700

TTCTGcTGCG CCATTTGCTC CGTATATCTT ACCGTTGCTT GCATCAAGCG CCGCTTGAAT   5760

TGCTGcGCGT TGCGCCTGAT GAAAGTTGAC CACCAGCGTA GTTTCTTCCT TTTTCCCCTG   5820

TGTGCATCGC ACCATGCAAG AGCTATCGCG TATGCACAGA GGTATAGCTG GTTGAACGCT   5880

TTCTGACACT TCGGTGTGCT CCCGAGGATA GACAGTAGTC TGCGTGCCGG ATTCCGGGGT   5940

CCTACATACA AGGTGTTCTG CGCCGATGGT ATCTCCGCGT GCAAGAAGTG CAGCACGCTC   6000

GAGTAGGTTG CGTAACTCAC GCACATTGCC AGGAAACGTG AGCGAGAAGA TTTTCTTAAA   6060

CGCGCTGGGA GAAAGCTGAG TGCGCTCAAA CCCCGGGCGG GTCTTAATTT TTTGGATAAA   6120

ATGCTCCGCT AGAAGCGCAA CGTCTTCTGC ACgCTCGCGC AAAGGGGGGA GACTGAGGGG   6180

AAAAACATCG AGCCGGTAGA GGAGGTCTTC CCTGAATTTC CCTTGGGTGA CTGCTTCTGA   6240

AAGGTTGATA TTCGTAGCTG CAATAATGCG AACCGAGACG CTTACCGAAC GCTCCCCTCC   6300

AACGCGCTCA AATACTCCGT CTTGGAGTAC GCGGAGGAGC TTCGGTTGCA GTTCCAGGGG   6360

GAGATCTCCG ACCTCATCGA GAAAAAGGGT GCCACCGTGA GCCAGTTCAA ATCTTCCCCG   6420

ATGGGTGCCG ACCGCACCTG AGAAGGCACC TTTTTCATGT CCGAATAATT CGCTTTCTGC   6480

AAGGCTATGG ACGAGTGCTG AGCAATTGAC GGGGACGAAG GGCTTGTCGC TGCGGGTGGA   6540
```

```
AAGTTGGTGA ACGGTTCGCG CAACAAGCTC CTTTCCAGTG CCGGTTTCTC CACAAACAAG   6600

GACAGGGAGG TCAGAGGCTG CTACGAGCTT TATAGCATCG AGTGTGCGTG TCCAAGCAGG   6660

AGAGGTTCCG ATCATATTTT TAAATGCAGG TGATTGGGGA GCTAAGAGCG CATTTCGTTC   6720

GGTCAAAAGA GCGTGACTCT TTTGACTCAG TGTCTCGGAC GCGTCGGTCT GGGCTACTGC   6780

GAGCGAGATA AGTTTAGAAA GAGTAGTAAT GAAGCGTACA ACGTCTGGGG TAAACTGCTC   6840

GCACAGGCGA TGGTCGAGCG TGAgCATGCC AATGGGAGTA TCATCGATGT AAAGCGGCGC   6900

GATGAGACAG GAATGATTTT GGGGCATGGG AATAAGCTCT GTGTAGGTAT CAGTGTGGGC   6960

AAGCGTCGGA TCGAAAAGGT ATGGACTCTT CTGTGATAGG ATGCGCGCAA GATCCTGCCT   7020

TTTGGTGAGG TCTATGgTGT GGTGTTGGAG GCGGGGAGTG TACAGGGGAC CACGCGCCTT   7080

GCGAACTCGC AGTATTTGAG AAGATTCAAA GCTGAGGACC ACGGCTAGCT CgTAACGGGC   7140

AATCTCATAG AGGcGTCCAG AATCATTtCC AGCGACTTTT CCGCAGcAGG GGGAGAGCGC   7200

GCGTGCAGGA CAGCCCGAAC AgTTCATGGG GCCCGAGTAT AGAAGAAAAA gGCATATCCG   7260

TGCAATTcTC CGCATGGAGC CTGTGGGCGT GTCGTGTGCA GGgGTATGGT ATTTGTTTTT   7320

CGAATCCTTT TCCtCGCGTT TTTTATGGGG TATAATCGCG CGCATGAGAC GCGTGTGGAT   7380

AAGTGTTCTG ATGTTTCCTT GCGTATGGGC AAATGCGCAG GGAGAATTTC TCGCAGGCGG   7440

yGCAAAGGGA TTGTACCGTA TTACTCCTTA CGCTCAAGAC GTACTGCTCT CTGGCGTTTC   7500

GGTTAGCAAG ATTATTGCTG CGGGAGAGAA CTGGTTCTTG CTTACGTCTC GAGGTGTCAT   7560

GACCTCGCGC GACTTAAGGA CTTTCGCGCA CGTGGGTGAG CAACTACCAA AGAAGGTAGT   7620

GAAGAAGATA GTCGATCGGG AAAAGGTTTT TGTGTCTCAG CCGCAGCCAT TGAAAGATCT   7680

TGAGGTACAT CCGGATAACG GAGCGGTTTT GGTTACCGCT ACCAATGACG CAGTGTTTCT   7740

CAGCAAAAAT GGGGGACGGA CTTGGCAAAA TCTGGGCTGT AATGCAAACA GCAGTGGGAT   7800

TAAGGCGGTc GCGGTGCTCG ATTTTCCTGA TGAAACGGGT AAGCCAGTGC TTACCGTGTT   7860

TGTTTCGCAT TCCCTGCGTG GTATTGCGTG GATGCAGCCA GAGAAAGGTC GTTTTTGGAC   7920

TGATATTTnAn GCcTwCnCTT GCGCTTGGTC CTGAAGCCAC TGAAGAAATC TCAGACATTG   7980

CGGTGCGCAG GAGCGTGCAT GGCAATGAGC TTTTTGCAAG CTACACGTTC GTGCCCAAGA   8040

TCGTACGCCT TAACTGGGCC AAAAAACGCT TTCAGGACGT ACGTGTGTGG AnCGntGCGC   8100

TGAAAGATGC GCGCTGCATT GATGGATTGA GTGCGTCTGn CnTTCGCTCG TTGGGTGTCG   8160

GGATGGTAGT TTGTTTGAGA TCCCCCTCAT TATGCCTCGC CCCTTCGATT TGGCGCGTCT   8220

TGAACAGGAT TTGCGTCGGA TCCCGGATCA AATCTTATGT GCGTGGGTTC CGCGTCATGT   8280
```

GTCACAGACG GGTGATGCGC TGTCTCTTTC TGAGTTATGG CTTTTGCACG ACCGGTCTAG 8340

TCTTGCAAAA GAGGGACGTT TTGCGCGGGC AGATTTGAAA AAGGGTATCT ACGTACCGGC 8400

GCATCACATT AAGGATCCTA AGCTTCGTGC AATGCACTTC AAAACGATTG CGGACAATAA 8460

ACTTAATATG CTTGTGTACG ACATGAAGGA TGAGCTTGGA ATGGTGCGTT ATCAGTCGCA 8520

AGATCCATTT GTGCGATCGG TCGGTGCAGT TCGTCCTTTT GTTGATATGA AGACGTTTGT 8580

GAAGCAGGCA AAGGAAAAAA AGTTGTACCT AATAGCACGT ATTGTGGTGT TTAAGGACAA 8640

GTATTTGTTC CGTTGGAATG GCTTTGAGCT TGCGGTTAAG GCGGGTGGTA AGCCTTGGCA 8700

GGGGTATAAA AACGGAGCGC TCCGTAAGGA AGAAATTCAC GAGCATTGGG TGGATCCGTA 8760

CAACGAAAAG GTGTGGCGGT ACAATGTCGC CATCGCGAAA GAAGCAATTG AGTTTGGCTT 8820

CGATGAGGTA CAGTTCGATT ATATTCGGTT CCCTACCGAC GGAGATAACC TTCACCAGGC 8880

GGAGTATCCG GCAAGAGAGT CCGGGATGGA TAGGGAGAGT GCGTTGATGT CGTTCCTGGC 8940

GTACGCGCGC GAGCaTaGAC GCGCCAATCT CCATTGATAT CTACGGAGCG AACGGGTGGT 9000

ACCGTACAGG CGCGCGCACG GGCCAGGACG TGGAACTGCT AGCTGAGTAT GTGGATGTGA 9060

TCTGCCCGAT GTTCTACCCC AGCCACTTTA GCCAGAGCTT CCTAGCTTAC GCGCCTGCGC 9120

AGGAGCGTCC CTATCGCATC TACTACTACG GCGGTACCGC AACCGGGTGC TGGCCCGCAA 9180

CCGGGTGGTC ATCCGACCCT GGGTGCAGGC CTTCTACtAC CGGTCTCTTA CGACCGGGCG 9240

TACTACGGCG AGGATTACGT GCAGCGCCAG GTGGCTGGCA TCCGCGAATC AATCGATGAA 9300

GGATACACGT ACTGGAACAA CTCAGGGCGT TACTCAGACG TCCGGCCCGA CGGCGCGCGC 9360

CTCCGTTAGC CGCCAAGGCA ACCCGGCGAG CGGACGCCCC TTGTTGGGCC GTTTTCCCAC 9420

ACAGGACCGA AAATCAGCGT CCGTCCTCTC AGGAGAAkmA CTTCTCTAGC TCGCCTATGT 9480

CCGCAGTCCT AAAAAGCTCG CTGCTGTCCT CGTCAAGGAT AACAACCATT GGAACGCTAC 9540

CTATACCGAA ATCCGAGACC AACTTGGCCC CTTCGGGGGG AAGTCGATTC AGCAATCTTT 9600

TTTTCTACAT CTGGCAGACC ACTGCAATAC TCGCGCAGGC GCGCGCTtCG CCTCGTCTGC 9660

CCCTACCACA GCGTTGACAA CCATAGCAAA CTCCCTCCGT TTCCGTCAGT CTGGGGCTTA 9720

CCACCCCCCC CCTGCACCGA TACCGCACAA GCTACGCGAG TTCAGTCAAC TCGGCGCACA 9780

GCACACGCAA AAGACCGCTA CTCGGTGAGG ATTTTAATAA CCTCCGCTTT ACTTGTTGCA 9840

TCAAGGATCT CCCGACGTTT CTCAGAACTC TTAAACAGGA GACTGATCTC AGCCAAGAAC 9900

TGCAGGTGCG GACCAGTGAC GTCAAGTGGA GAAAGGGTCA TTATGAAAAT ACGACAAGGT 9960

TCTTGATCCA AAGAGTCGAA GTCAACCGGA CTGTCAGAAA CACCAACCCC TGCAACCAGA 10020

```
CTACTCACTG AATTAGTTTT ACCGTGGGGT ATGGCAATGC CATGCTTCAT CCCCGTAGAC    10080

ATTTTTCGCT CTCGATCGAG CACACACTCC CGCGCAgcAA CCTTGTCGCT CACCCTTCCT    10140

GCACGGACGA ACATCTCGAG CATTTCGTCG ATGATCTCCT CCTTGGTAGA ACCCTTCAGG    10200

TGCAGGCTTA CGGTTTCCGG CGTCAACACG GTCTCCAAAT TCATTCCCCC AAGCTAAAAA    10260

CTCTCAAAGG AAAAGTCAAG CTTTTTGGAA AAGCTCCCCA CGCTCTTTCC GCTCCCAGGA    10320

TATCTTGACC TTTTGCCTAC TTGGACTTTA CCATGCGCGC GTGGAGTTCG CCACCAGGGA    10380

GCAGCTGAAT AGGTACTACG ATTTGTACAA GGATGTCGAT GTAACTTTCT CAAAGGATGT    10440

GATGCAGGCG CTCTGTTTTA ATGCGCGGCA GGTGTGCGTG CGAACGCGGn GAGGTCAGTG    10500

TTCCTGCGTA ATGAATTCTG TGTCTATGGT GGGTGCGAAA GTTATTCTCA GCAGGAAGAG    10560

TAGTCTGCTT GAGAGCATTC AAGTGGAAGG GGCGAGCGTC AgCATACGGT TTTCTTTCTT    10620

TGAGTCCGAT GCGCGGGATG CGGTTTCCTT CTTCGTTACT GCCAGGGTTC TCGGTGTTGA    10680

AGACTATGCC CaGAGTACGG AGCTAGTGGT GTTAAGCGTG GCGTATACGC AGCGCATACC    10740

TGATATGCTC ATAGAGCGTT TGGGTTTGCT TGTTGAGGCC AACATTAGTT CCAAGAAGCG    10800

TAAGTCGGAG CGTATTGCGg TGAACAAGGA GAGTATGCGC AGGATCGGCT TGATGAGAGC    10860

GGAGACCATC GTGTTCATTC AGGCGATTCC TCGCCGCTGC GTTCTGCGGG ATGTTTCCTT    10920

TGGTGGTGCG AAGTTTATCA TGATGGGCGT TGCGCCGTTT TTGAAAGGCA AGGAGACGGT    10980

GCTGAAGCTT GATTTTGAGG AGCCGAGTAC GAGCATGAGT ATTAGGGGGC ACGTGGTGCG    11040

TGCAGATCAG GTTGAGGGGC GTAAAGACCT GGTGGCCGTG GCCATGGAGT ACGACTTTGA    11100

TGTGGTGCCT GTCGCGTATC GTATGTGTTT GAACCgsTAC GCATCGGACC gCTGTCGCCG    11160

TTTTCCCGGT ACGGACGAGG ACTGCTCTGC GGCGTCTGCC GGCGATCCAG GGCGGTCGTC    11220

AGCAGGCGCT GAAGGTATTG ACCTTTCTGT ACCCTTCTCT TTGTCTTAGT TTTAATGGCG    11280

CTTGTCACCG GATTGCCCCT TAGGGGGGTC CGCATGTCTG CGGATGCGCG CGAGGCTCCG    11340

TGCACTGAGC CTCGCTCCGA CCAGTAGTTT CGAACGCACA AACCAGCCTC GTgGCACTTC    11400

CAGTGCGTAC CGTACTGACC GGTTGCTGCG TACGCTGTGT GTGCTCAGCG GGACTAAGTC    11460

TTCTATCTGT ACGATCGCCC CTTGAGCATC CAGGAATGCG AGAGAGAgCG GGTGTGGGGT    11520

GTCCTTCATC CAAAaGGAGA GGCGTGTGTC CTGTTTATAC ACGAAAaGCA TGCCgTCCcG    11580

TCGGGGATCc GTGTACGCCC CATGTACtCg CGnCCTGCGC TTCTTCCGTG AGTGCGAGTT    11640

CTACAACCAC CGGCACGTAC TGCCCTCCTG TACAAAAAGC GATTTGCGCT GTTTCTAGGC    11700

TATTCGTTCT GCACGCCACA CAGGAAAGCA ACCCCAGTAA CAAAAGCGAC AGCGCAgCAG    11760
```

```
GTGTCCTGTG CAGGTAGAAG GAGACGTGCT CTTCAAAAGC CTTGTACAAA ACGCTTTCGA   11820

TCCTTTTCAG CATTGCTCTT TTGAGTGTCG CGTGCACTAA GCTGTTCGTT AAAGGCGCGG   11880

ATATCTATGT ACTTGATAAT AAGAGGCCGC TCGAGCACTA GGCGTGTGGA CGCGTCTTCC   11940

CACAAGGCAA CCCGAGGACT GAGCTTATGA GGTTCGCCGT ATTTTTTACA CAAGTTCTCA   12000

TATACCGAGT AGTAATCTAT CGCGTCAGTG TTGAGTTTGA ACGTCATAGC ATATAGACGG   12060

TCACGGTAGA ACTGAAACCA GCTGCGTGCA ATAAAGTGGG GACCCGTGGT CTCAATCAAG   12120

ATACGGTTCT CACTCATTAG CAGTGACACG TCACGCTCGC CGCGATATCC AAAAATACTA   12180

TCCTTTTTCA GTGCTTCCTT TACCTCGGTT ACGCCCATAC CCAAACTCAG CGCGCGGTAC   12240

ACGGCGGGAA TTTTCTCAAG AGAATGTGTC TGCGAAAGAG GAGTGACTGT AGCCAGCATT   12300

CTCGCACCGC CTGCGCCGCG ACTAGGGGAA GGGGAAGAAC ATAAATAAGA CTGCATATGC   12360

AACACCCCGC CCAACGCAAG CGCCATACAA TTTTTAGCAT AATTCATGCC GTTTTCCTCT   12420

CCTTGTGTGG ATACATCTAC TGCCTGTTGT GTGGGTCCTT GCGTGCCTTG GCGAGGGTAC   12480

GCTTTTTCTT AAAATAGGCA ACGCGCTTCA CCAATTCTAG GTCGTCGGTG ACAGATATTT   12540

TGTTATCGAC ACAGCGGACA ATAGGTTCTC GTAAAAAgTC TtCGGTTACC TGAGAGACGA   12600

TGTCTTTAGG GAAACCGCAC ATGTGTGCGA GTTCTATGGG ACCGAATTCA AAATCGTAGG   12660

CCTTGCCGGT GCTAGGTATG TAGCGAATCT TTTCTAACTG GATGGCCAAC ATATCGTACA   12720

TTTTTTCAGT CGGTTCCGGA AGCAAAGTAT TTGCGAGCTG TCGGTACATC GACCAGATGC   12780

GATCTGCGAG CGTGGTGGTA AGACGCGCAg TCAATTGCGG TTGTGTGGCT ACCAGCTGTT   12840

GGAAGTTCTT TCGGTTCACG GCCAAAAGCT GGCAACCATC AGACATAACA ATGGCGCTTG   12900

CAGAACGCGG CTTGTTCTCC AGCAACGCCA TTTCCCCAAA CATATCTCCT TCTTTTAAAA   12960

TCGCCAGCAC TACCTCATTG TTATCAACAA TCTTAGTAAT TTTTACATGT CCTTTTTGAA   13020

TGATGTAAAA CTCATTTCCC AATTGACACT CACAGAACAC CATCGCCTCT CGATCGTAGC   13080

AGCGCGTGGC TTCAAGTATG TTAGGTTCGA GTATTTCTAC TGGTACCTTA ACTCCTGTGG   13140

ATTTAATCGC AACAAATCGT TTGCGTGCTT CCTCTGCATA CGTCCCCTTG GGACTTTCCT   13200

TGAGATAATG ATAGTACGCA TAGAGCGCAA GTTCAAACTT CGTCATTTTG ACGTAGTATT   13260

CGCCAATAGC GAAAAGATGC GAGACATCCA CATCAGTGTG TTTTTTCAAT GTCAATTGGG   13320

TAAGCGCCTC ATTGAGGTAG CGCATTTTCT TTGTGAAGGA AAGGATGATT TTCATAGCAA   13380

TCGCCGCGTT CTTTTCAATG AGCTGGGGGA ACTGCTCATA ACGAATTGCA ATAAGCACGA   13440

CATCAGTGAG CGCAACTGCA GTTTCAATCT GATTATGCCG CGACATGCAG GCAACTACAC   13500
```

```
CTAAAAAGTT ACCCGCGGTT AGGACGTTTC CTTCCTCCTC TGCAACTATC TCTACTTGTT    13560

TTGCAATACA TACCTGACCG CTGTGAATGA TATAGAAAAG ATCTGCGTCA gcTTTCCCCT    13620

CTACTAGTAT GTAAGAACCC TTCTTGAAGT TAACAAACGT CAGCTGTAAC AAAGTATCCC    13680

ACTCCTTCCT AATCGTCGCT CTATGCTCTG TTTACAAGAC AACCCGTACG TCTTGCAGTG    13740

CACAGGTGCC CGCCTCATGA GTCTCGCGAA CTGGAAGTCA TCCTCACAGA TTCTCCAGAA    13800

AAAATGATTT CCCGTTCGAG CCACACTCCG TGTGTTTCAA ACACGCGTTG CCGAACGACG    13860

CGCAAGAGTG TGCGCACCTG ATGTGCGGTG GCATTCCCCG TATTGATAAT GAGATTTCCA    13920

TGCCAGGGTG CTACCTGCGC AGcCCCACAG GAGGTGCCCC GTAAACCTGC CTCTTCTATG    13980

AGAATGCCAG ACGGTTTACC AAAAGCTGGG TTGTTTTTAA ACGCGCTGCC TGCTGACGGA    14040

AAGCGAAACT GCCCCTTTGA AATACGATCG GCAATCTTCT CCTGCATGTG CTTCCTAATC    14100

TGCGCCGGAT TGCCGGGAGT GAGACGTACA CACAGCGAGA GGATAAGACG CCTTCCTGCA    14160

TGGAGTTCAA CACCGTGAGG ACTCTGGAAA GGAGAGCGCT TGTAGCCCCA ATCCCCGCGC    14220

GCGCGAAGAC GGTCTGAAAA CTGCTGCAGG TAAACGGTCC GCCGTCGAGG CCAAGACATT    14280

CCCCTCTTTT GTCTTGTGCG TTTTTTCTCA CCTCTGGCAG TTCTTTTGCG CGCGAACGcA    14340

CGGGGTGAAG TaCGAGCGTG CGCGCAGAGT GAAAnCAAtC TGCGATTGCA CGCCCATAAC    14400

ATCGGGCGTT CATGTACGCG GCACCACCGA CACTACCAGG CAGCCCTGCA AAGGTCTCAA    14460

GCCCGCGTAG AGCGTGATGG GCACAAAAGG CCAGGAGGGC GGCCACAGGT AACCCCGCGC    14520

CTGCATGTAC GAGCACTGAG CCATCGCGCT GTGTTTGGGT GTGTAGACTG CGAAAgCGAC    14580

GAAGGCTCAA CATcAGACCC GGTACGCCCT cGTCTGCGAT TAACACGTTA GAGCCTCCCC    14640

CAATAAGGGA CAGCGGAATG CGTGCGCgCT GCGCTTcCTC AATAAGCGCG CGCAgcTGTG    14700

TGCAGGAGCG CGgcTCCGCC CAAAACTGCG CAGCgCaCCA ATGCGGAAAG AACATCGCTC    14760

TGCAAGTGGG ACGTTACGGC GCGTGATCCG ACGCGCGCGT ATCCGGTGCG CGGACATGGA    14820

CAGAAAACAT ATACGATTTC ACGCGCTTAT GCTACAATGG CGCGGTCTTG GCCTTCTTTG    14880

CGTTTCGAGG GAGGGTAGAC TGAAGCGCAG GCGGGCAAAG GCGTTTGCGC GACAATGGGG    14940

CTGGGCGTGG TCCGCACGTG TTCTGCGCCC GGTTGGGAGA AACTTCGAAG GGGCGCACAT    15000

GGCCTTGCGC GTATACAACA CCCTTACTCG TCAGCAAGAG CACTTTCAAC CCTGGGAGCA    15060

CGGGCACGTG CGTgCTCTAC GGTTGTGGGC CTACGGTGTA CAATTATCCC CATCTGGGGA    15120

ATCTGCGCGC ATACGTTTTT CAGGATACGG TTCGACGTAC CTTGCACTTT CTTGGATACC    15180

GCGTCACCTA CGTTATGAAT ATTACCGACG TTGGGCATTT AGAAAGTGAC GCAGACAGTG    15240
```

GTGAGGATAA GCTGGTAAGG AGCGCACAGG CGCATGGCCA CTCGGTGTTG CAGGTTGCAG   15300

CGCACTATCG CGCAcCTTTT TCCGCGATAC TGCACTGCTC GGTATTGAAG AGCCGTCCAT   15360

TGTCTGTAAT GCCaGCGATT GTATCCAGGA TATGATCGCG TTTATCGAGC AATTGCTCGC   15420

GCGTGGGCAC GCGTACTGTG CAGGAGGGAA CGTGTATTTT GATGTGCGAT CCTTTCCTAG   15480

CTACGAAAGC TTCGGTTCTG CCGCGGTAGA AGATGTTCAG GAAGGAGAGG ATGCGGCGCG   15540

CGCGCGGtGG CACACGATAC GCATAAGCtG ATGCACGTGA TTTTGTGCtG TGGTTTACCC   15600

GTAGTAAATT tGTGCGTCAT GCGTTGACGT GGGATtCTCC GTGGGGGCGG GGGTACCCCG   15660

GTGGCACATC GGGTGTTCTG CAATGAGCAT GAAGTTTTTA GGACCACGTT GCGACATCCA   15720

CATCGGAGGG GTGGATCATA TTCGTGTGCA TCACCGTAAC GAGCGTGCTC AGTGTGAAGC   15780

AATTACTGGT GCACCCTGGG TGAGGTACTG GTTACACCAC GAGTTCTTGC TGATGCAGCT   15840

GCAAAAGCGC GCAGTACATG CGGATATGGG CAGTTCGgTG GTGTCGTCTT TTTCTAAAAT   15900

GTCCAAGTCC TGTGGGCAGT TTTTGACGCT TTCTTCGCTG CAGGAgCGTG cTTTCAGCCA   15960

GCTGATTTTC GCTTCTTTTT GTTGAGTGGA CAGTATCGCA CGCAACTTGC TTTTTCTTGG   16020

GATGCGCTAA AAACGGCGCG TGCCGCCCGA CGGAGTTTTG TGCGGCGAGT GGCGCGTGTA   16080

GTGGACGCTG CTCGAGCAAC TACAGGCAGC GTGCGCGGCA CTAGTGCAGA GTGTGCCGCA   16140

GAAAGGGTGT GTGAATCGCG CGCATCAGAA TCTGAGCTGC TCTTAACTGA CTTTCGTGCT   16200

GCGTTGGAGG ATGACTTTTC TACGCCACGT GCTCTGAGCG CCTTACAAAA ATTGGTGCGT   16260

GATACCTCGG TGCCGCCATC GCTGTGTGTT TCGGCACTCC AGGTGGCGGA TACAGTGCTA   16320

GGGTTAGGCA TAATACAGGA AGCGACCGCA TCGCTATCTG CGCAGGTTCC TGCTGGCGAT   16380

ACGTTGCCGC AGCGTCCTTT ACCGAGTGAG GAGTGGATTG GACAGTTGGT GCGTGCGCGT   16440

GCACATGCAC GCCAAACGCG TGATTTTCCC CGTGCAGATG AGATCCGTCG GCAGTTGAAG   16500

GCTGAAGGGA TTGAACTTGA AGACACCCAT CTTGGGACTA TTTGGAAGCG CGTGTAACAT   16560

TTTGGGAGAT ACATTGTTGC ATGAGCAGGA GCTTTTAAGA GCACAGGATG ATGCAGATTT   16620

TAAGCTCATG TACGAGCAGC TTGTGCCAGT GCTCTAsCGC GTAGctAcAA CGTGGTGCGC   16680

GAGGAGGACA TCGCTGAGGG GCTCTGCCaT GATGCCTTCA TTGCAtGACA GAAAAGAGGA   16740

TGGAgTTTCC GTCTCTGTCG GACGCAAAGT ATTGGTTGAT CCGCGTGGTG AAAAATGCCT   16800

CGTTAAATTA CGCTAAGCGT CGTGTACGTG AGCGTCATTC TtGTGAGCAA GCGTCGCGCG   16860

AGCATGTGTG CGAGCCGGAT ACCGGTgrmT TCGCTTGTTA AGAATAGAGA CGATTGAGCA   16920

GGTGCGCGCG GCCTTAGATC GACTGCCCGA GCACCTCCGT GTGGTTTTGC AGTTGCGCGA   16980

```
GTATGGGGAC TTAAACTACA AGGAGATCGG ACGTATCCTG GGCATCAGCG AGGGGAATGT   17040

AAAGGTGAGG GTGTTCAGAG CGCGCGAACG ATTAGCGAAG TATTTAGGAG AGACGGATGC   17100

GTACCTGTCC TGATTGTGCT GCTTGGTGTG CTTATGTGGA CGGAGAAGGT TCGCAACTGC   17160

AACGCCGTGA GATGTGCGCG CATCTGCAGG GTTGCACACA CTGTGCCACG TGTGTGGCGC   17220

ACTATCGCGC CATGCGGAGT CTTGTCAAGC ATGCTGATCG CGTTTCTTCC CGTGATTTTA   17280

CAATGGCTTT TCCATATTTG CGCGTGCGTC ACCGTGTCGC TTCCTGTATG CCGAGGCCGT   17340

GGTGGCAGGC ACGTTCCTCT CCTCTTTCTG CTGCAGGACC GGTCCgTGCT GCGGCACTCG   17400

CTGTGGCGGT CGCATCTTTA TGTGTATGCA CCCTGTTGCT TACTCATATT GTTGAAAGGC   17460

GTCCTGTATC CCGTGCGGGT GAGGCGAGTT TTACCCCCAT TGTACCTATG CGTGTTCGCG   17520

CCCCTGTTGG GTACGCGCGC GGTGTGAAAG TGTTTGGTCC TGCCGTTAGT GCGAATTCCA   17580

ACGTgTGCGC AAACCAGCTG CGGTGTTCAC CGTCTGTGCG TTTGCGCAGT TGTATGGCTC   17640

AGATCCTGCG TATGAAATGG AAACAGTGCC GGTGAGGCTA TCGGTTATTC CTGTGCCTTC   17700

CTATGTGCTC AATGCTTCAA AAGCGCAGTT CTTTTCCCCA TAATCCAGGC AAATGTGTAG   17760

TAAAAATAAT GCGCCCGCGC GGACGTGTTT CCTGTTCTTT TCAAACCGTT CTGAtCGTTG   17820

GGTGTTCCTG TCTGCAAACT TGATTGACCT GCTTGTCAGG TAGCCATAAG GAGAATGTCT   17880

ATGACCTTCG TTGAATCAAT GCAGCGGCGT GCTGTGcTTG CGCAAAAACG ACTCGTGCTT   17940

CCTGAGGCCT GCGAGCAGCG TACGCTCGAA GCCGCCCGTT TGATTGTGTT CAGAAACATA   18000

GCCGCAAAAG TTTTTCTTGT CGGATGCGAG CGTGATATCA AAAACACCGC AGACAGGTGC   18060

GGTATCGACC TTACCGACAT GGTCGTCATC GATCCGAGCG TTAgCAAGCA CAGAGATCAG   18120

TTCGCAGAAC GTTATTTTCA GAAGCGAAAA CACAAAGGAA TAAGTCTTGC CCAGGCTGCA   18180

GAGGATATGC GCGATCCTCT GCGTTTCGCT GCTATGATGC TTGACCAAGG TCACGCAGAT   18240

GCCATGGTTG CCGGTGCAGA AAACACTACC GCGCGCGTTC TTCGTGCAGG CCTCACCATC   18300

ATCGGAACCC TTCCGAGTGT TAAAACTGCC TCTTCCTGCT TCGTTATGGA TACTAATAAC   18360

CCCCGTCTGG GAGGAACACG TGGTCTATTT ATTTTTTCAG ACTGTGCAGT GATCCCCACT   18420

CCCACCGCAG AACAGTTGGC TGATATCGCC TGCTCTGCTG CAGAAAGCTG CCGCACCTTC   18480

ATTGGAGAGG AACCGACTGT CGCACTTCTT TCCTACTCTA CTAAAGGATC AGGAGGTGAT   18540

AGTGACGAGA ATATCCTGCG TGTACGTGAG GCAGTCAGGA TTCTACACGA ACGGCGGGTG   18600

GACTTTACCT TCGATGGGGA ATTGCAGCTC GATgcTGCGC TCGTACCTAA GATTACCGAA   18660

AAAAAAGCGC CTCACAGTCC TATTACGGGA AAGGTGAACA CACTCGTGTT TCCCGATCTT   18720
```

```
TCTTCGGGTA ATATTGGGTA CAAGCTTGTC CAGCGCCTTT CAGATGCGGA TGCATACGGA   18780

CCTTTCCTGC AAGtTTTGCA AAACCACTGT CTGATCTCTC GCGTGGGTGC TCGGTTGAAG   18840

ATATCGTCGC CGCTTGTGCA GTCACACTTG TGCAATCGAA TGGACGCTAA TGACGTCCAC   18900

CCAGGCGCGT ATACGTGAGG CAGTCCGTGC AGGGAGCGTC CGAGATTATG CGCGTGCTAT   18960

CCGTATTCTT GAAGAGCTTG CCGCTTCAGG AAAGGCAGAA GGATGTCATC ACCCAGATGG   19020

CGGTGCGGTG TATGAGAGGG GGGCACAGGA AGAGTGGAAT GAGGGGTCGT CTGAGTCGCA   19080

CGCGCACGGT GGGGATGGTA CGCAGGACGC GTATCCTGAG ATTTATTTGT ATCTTGCGCG   19140

TGCATACCAC GCACAAAGGC AGTATGCGCG CGCGGTAgTA ACGCTACTGT GTATTCTAGG   19200

CGCGTGCCGC gcGrACGGCG CAGGTTGGTT CTTTTTGGGA AGGAGCTATC TTGCACTGCA   19260

TCAGGGGGGG TATGCGGTTG CAGCGCTTCG GCGCAGTGTA CGAGAAAATC CTGCCTCTCT   19320

TGGGGCGCAG GCGCTGTTAG GACTCGCCTA TCTGCGGAGT AAGAAGCCGC GTGCAGCGCG   19380

CATGGTGTTT GAGCAAGCAC TTGCGCAGTA TCCAGACAAT AAGCGTTTGA ACGCAGGGTA   19440

TTTGAATTCG CTTTTTGTAG AAGCAGTGCA GCATCTAAAA CGGGGGAGCG CAGATCTTGC   19500

GCGTCAGATG TTTACGTTTC TGATTAATCA GGATGTAGAC GGGGTTGCGC CACGTTTATA   19560

cTTGGCGCAC GCGTTTCGTT CTTTGAAACA TTTTCCTGAA GCGCTTACCC AGTATCGTGC   19620

AGCAAGCGCA TTTGCGCCGC ACGATCCTGC CCTCAAGTGG TACGAAGCGG CCATGCTTGT   19680

AGAAATGGGG TGTCTGTCGC AGGCGGCAGC GTTGCTGTCG ACGTTGGGTG TTTCCATCGA   19740

GCGTGATCAG ATTTCGGATC GTTTTCTAGT GATGGGCGCC GTGCGCAAGC ACATGGAGGA   19800

GGGGGCGTGG GCTCGTGCCG CTTCTGCAGC GCATTTATAC CTGAAAACTT TTGGGGGTTC   19860

TGTAGAAATT CACCTGCTAA TGGCAGAGGT TCACCGGCGT GCGGGGCGCG TGAACGTGGC   19920

TTTGAACCAC TACACGCGTG CGATGAAAAT AGAACCGAAA AATTGTTATC CGCATTATGG   19980

TCTTATGGTG TGTTTGCAGG AAGCGAGGCG CTGGCAAGAG CTGGCAAAGG CAATCAGACG   20040

TGCAGAAGGC GCAGGGTGCG ACGCGCAGGA TTGCTACTAC TACCGGGTGA TTACAGCTGC   20100

CCATTTGAGC AATCtCCCGA GGAGGTGTTA CCGCATCTGC AAGAACTTGC GCGTGGAGGG   20160

AAGGCCGATC AGCTTTTGTT CAATGCTCTT GGGGTAACGT ATGTGCGACT GGGAATGGCA   20220

GATCTCGCAc TTCGCTGGTA TGAAAAAACC CTTCTTCTGG ATGCAGAGGA CGAAGAAGCG   20280

TGCGTGGGAC TGATCGCCTG CTAcGAgGCG CTCTGCGACG AagCGcGCGC GTACACCCAG   20340

TATGGAGCGT ACCTGTCCCG CTGGAGGGAC AATCGGGTTA TCCGcAAGGA TTTTATAGCC   20400

TTTCTTGAGA GAACAGAACG GTGGTCcGAA GCGGCGGACC ACATCGAGTT GCTCGCCTCG   20460
```

GGTGAGCGAG GGGGTTTTTG GGGTACTCGC CTTGCGTTTG CGCGTAAAAA AGCCGGCCAG   20520

TACAGGCAGG CTGCAATTAT CTACCGGGCG CTCTTACGTC AGAGACCGGA CGAGCGGGTT   20580

TTACTGCACA ACTTGGTATA CTGTCTTGAC AAGATGGGGC AGGCAGACGC AGGGCTAAGG   20640

CTGTTCCGCG CTGCGTGCAA CGCGTTTGGG ACGAGCGTGG AA   20682

(2) INFORMATION FOR SEQ ID NO: 71:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 1356 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: double
       (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

TTTATGCACC CCAnTGAATC GACAGCCCGA CTTCAGAnCA CACAnCCCGC GCAgCCACAg   60

GATGAGCTcC TGCGGCAAaT GGTTACACAA ACACCGTCAC TTCACCGCAG CATATATCCC   120

TGCACCAATm rCGGCTACAC CGCACACAAC TGCAACCCCG ATAAGAACAT TATTCGTGAT   180

CTCTAAACGC CTCAACCGCA TGTTCAAykm GTTCACACGC TGCCTCAATA TCTCCAATTC   240

GCTTCTCAAT GTCTCTATCA ACTCTTTCGA TTCGCTCAAT GCGTGCTCCG ATTCCTcCAA   300

GgCCTTGGCG GCCTTCTCCA ATTTCACGTC GAGCGTCTGT AAGGCCTTTT TGAGCGCGGC   360

CGATTCGGCG TGCCGTTCCC TCAACTGCTG CTTGAGCATG TTTGATTCGA GCCGGATCGA   420

TGCCACcTCC CCCATAATTT CCTTTAAAAG CCCACCAGTA GCCGCCTGCG AATCCGCATA   480

TGCCACAAAA GAGCGCAACA ACACCATACC CCACAATAGT GCGCCCACAC CCCGCTTCCA   540

CATTCGGTCT CCTCACGACG ATGCGTTCAC CTTCCATTCA TGAATAAATC CAAGCATGTA   600

TTGCTGGATA TCCTCAAACC ACTTCTTTTG GAGTGCGCAC GCAGCaCCCC TACATAACGG   660

AAATGCCACG GCTCCCTAC ATACCCCGTC ACCTGCTCGT AACCAGGGGG AAAAGACAGC   720

GACCATCCAA AACGATGGGC GTTGCGCTGC GTCCACCTCC CTGCATCACT CCGTGCAAAC   780

GCCGGCGTGA TAGAACCGAA ATCCACTACC GTCCCCAACT GGTGCTGACT TGTTCCTTCT   840

CGCGCGGAAA AACGCATAGC CTCCTGCATG CCATGCTCCT GCGCATACCA GGAGAACAAC   900

TTTTTCTGAT ACGCAAAAGA GCGATAGGCA GAACCAACGG ACAGTGCCAC CCCGTCACGC   960

GCAGnCGCCT GAATCAGCTG ATGTAACGCT TCGTACGCAA TCTTAGTTAA AAGGAGCGAC   1020

CTCCCCTTTG AAAAAAGAG CCACTGCTCA CGCACCGGCA CCAGATGCTT CGGCACGAAC   1080

GTTTCTGGCA GGGGATGTTT CTTGTCAACC AAACGCAACA GATACCCCTC CGTGGTAAGT   1140

ACCGCGTCGA GCTCTTGCAA AAACTCCCTC CCCTGTGCAC ACAAACGCGT ACAAGGCGCG   1200

CAGGAAGCGC CGCGCGtTCG CAGCAGCACG CACACGATGC AGATCCACCC GATCCACGCC   1260

CTGcGGcGAG ACCGCATTTC CCAGAGCAAC TAACACGTAC CCTGGCATAC GCACACGCAT   1320

TCCAACGCGC CAnTTAGTCC AACTCATCTA ATGATT   1356

(2) INFORMATION FOR SEQ ID NO: 72:

       (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 4579 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: double
          (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

TACTGGTGGT ATCCACCTCA ATGAACTGTT TTTCCCTTTG TGGGAATCTT CTGAGGATCT   60

CCaTCAACGC GCAGGAGCTT GGCTGTATAC TTATGCACGT GATGCAGAAA AGGAGCTCTG   120

ACATACAATC CAGCGGTGTT ATTCGTTTTT ATGATTTCTC CGAACTGGGT GATGAgCGCA   180

AcCTGTCCTT CCTGGATGAG GTAAAACGGT TGCAGGwGCA CAACACCACC TAACAGCACC   240

CCAACGACTA TACCTATGTT CAGAACAGGT CGTAaCGTGC GTGTACCTGT AGTCCACGTT   300

TCCTCATACC CCTACTCCTC GCGTGTTCCT GCCACGACCT TCTTCGATAC CTTACTGATA   360

TCCTTGAGCG TTAAAAGATT CTCCAGTTTT TTGTCAATCA ACAGCACATT TTCAGTCTTT   420

TCCAGGATAG CCCCCAGTCC CTCAAGGTAC AAACGCGTTT TGGTAACATG AGGTGCTTTG   480

ACATATTCAG CATAGATTGA GTCAAAACGT GCTACATCTC CTTTTGCTCT ATTTACGCGT   540

TCATTCGCAT ATCCCATAGC CTCCTGAATC AACTTGTCCG CGTCACCTCG GGCCTTAGGA   600

ATTTCCCTAT TGTAGGACTC TTTTCCCTCG TTAATGAGTC GATTCATATC CTGAATAGCA   660

ATATTCACGT CTTCAAACGC TTGCTGTACC TCCTGAGGAG GAACAACATT TTGCAGCTGC   720

ACGGAGGAAA CAAGAACACC TAGGCCAATC CTTTTCAGGA GAACATTCAT CATATCCTTC   780

GCACGCATCT GAATCGCACT GCGCTCCGGC CCCATGATAT CAAGAATCGC TCGATCTCCA   840

ATTAAACTGT TCACCACTGC TTTTGAAATG TCTCGAATGG TTTGCCTTCG CTCCTGGGAC   900

TCAACATTAA ACACCCATGC TCTTGGATCT ACAATGCGAT ACTGAACCAC CCACTCGACG   960

TCTACAATAT TCAAATCCCC CGTAAGCATA AGAGACTCGT GACTGATATT ATTCACATAG   1020

TGACTCTGCT CGGAACTCTT CGACGTTCTG AACCCGAACT CTTCCTTTTG CACCTTGGTT   1080

ACCGGCACTT TATACACCCA CTCTACAAAG GGGATAAGAT AATGCAATCC CGGTTCTAGC   1140

GTCCGATGAT ACTTGCCAAA ACGGGTGACC ACCCCATTAT CAGTGGGAGA AATGATCCTA   1200

ATAGGGGAGG CAATTCCAAC AATCACGATA CCGAGCACCC CACCTATGCA TCCTGCCACC   1260

ACGCTCCACG TTGCTGGAGT CCACTTTGGT ATTCGCATCA CGGCACCTTC CTACACACGC   1320

TCGTCTTTTC GCCCATCTTA CGGGAAACAT TTTCCTGTGA CAACACTCAC CGTATTCACA   1380

CAGAcTTCGT TGTAGACAGA ATAAAAATTC TCACTCAGTA TAAAAACACA GGAGGCATGA   1440

TGTATCTTAC AAAGGAACTA CTCGATACGT TTGCGCACGA AGTCGCCGCA GATCCTATAC   1500

ACAAAGCGGT CGCAGGAGCT GTTGCGCGCG TCGGTCTTGA AGAAGCTGCA CTGAACACAG   1560

AAGTGGCGCG TCAgCACACA CATATTTTTT CTACCGAGAC AAAACGTGGA GAAATGACCA   1620

ATCAAAAAAT GAGTGGTCGC TGCTGGATAT TTGcTGCGCT CAACGCCGCG CGTGTAAACA   1680

CCATGAAAAA GTTGGACATT GAAACAGTTG AGTTTTCCCA AAACTATCTT TTCTTTTGGG   1740

ATAAATTGGA GAAAGCAAAT TTCTTTTTAG AAAATATCCT AGAAACACTT GATGAACCTC   1800

TCACCAGTCG GTTGATGGCA CACCTGCTTG CAAATCCCGT CCAAGATGGC GGGCAATGGG   1860

ATATGTTTTC AGGGTTATTA GAAAAATACG GTCTTGTGCC CAAAGAATGT ATGCCTGAAA   1920

CTTTTCACTC TTCCAACTCA CGCGTTCTTC TTGCAGTCCT CACTCGTCGG CTGAGGAAGC   1980

ATGCACAGCT TTTACGTTCT GCGCATGAAG AAGGCGTTGC GCTGCATACC CTGAGGGAGA   2040

AAAAGGAAGC GTTCCTTTCT TCCATCTACT CTATCCTCGT GAAGGCTCTC GGGAGACCTC   2100

CGGAGAAATT CGACTTTGTG TACAAGGATA AGGAAAAAAA ATTTCACAAA GTCAGAGACC   2160

TTACGCCGCA GAAGTTTTTT TGCGATTTCG TCGGATGGGA TCTTAAAAAC AAAGTGAGTT   2220

TGATTCACGC GCCAACTGCG GATAAACCGT TTGGCAGAGC ATACACGGTT AAATTTCTAG   2280

GCACCGTAAA GGAAGCCCCG TGCATCTGCT ATGTCAATAC TCCCATTGAA GTGCTCAAAG   2340

AAGCTACAGC TTCTGCAATC CGAGCCGGGG AGCCGGTATG GTTTGGTTGT GATGTAGGTC   2400

AAATGATGAC GCGCAAAGAT GGTATCATGG ATACGGAGAT ATTCGGGTAC GAGTCGATGC   2460

TCGGCACTAC CCCTGAATTC AATAAAGCAG AACGGCTTGA CTATGGCGAA AGTCTTTTAA   2520

CACACGCGAT GGTCATAACC GGTTTTGACG AGGATGCACA AGGTAACCCC GTACGCTGGC   2580

AGGTAGAAAA TTCGTGGGGA GATGACACAG GAAAAAAGGG CATGTTCTCT ATGAGCGATC   2640

GCTGGTTTGA CGAATATCTC TACCAAATTA CGATCGACAA GAAGTTCGTA CCACAGGTGT   2700

GGCTCGATGC GCTAGAGAAG CCAATAATAG CGCTCGAACC TTGGGATCCG ATGGGAGCGC   2760

TGGCGGACAC CCCTCTGTAT CTTAAAAATT AAGAAGAAGA ACAAGTGCGC AATTCTGATC   2820

GGTACTTATT TACGGTACGT CTTGCGCACT TGATGCCCTG CTCACCGAGC AACTGGGCTA   2880

TCcTTCGGTC GGAAAGGGAT ATACGCTGCG TTCGTACCTC TTGTATGAGA CGTGATATCC   2940

GGTACTTAAC TGATACTTTT GAGTGCGGAG AACTCGGGTA ATTCTGTCCA AGACTGGACC   3000

GATCACGATA TTCTTCGGTG GATAAAACCC GAGGGGAGAA AAAGTACCTT AAGGAAAAGT   3060

GTTGCGATCC GTACTGGAGC CATTTGTCGC GCACTATGCG GGACACTGTT GAAACGCTCA   3120

ATCCGGTCCT GTGTGCAACA TCTGTCATTC TCAGGGGCGT GAGctTCGCA GGTCCGTGAT   3180

CAAAGAAACC GCATTGGTAG TGAACTATTG TTTTCGCGAT ATCCAGCAAG GTACGTTCCC   3240

GGTATGAGAG CATACTTACA AGACTGAGCG CGTCGTGCAT GCATGCTTTC AACGCGTGGT   3300

TTTTTTCTGC CGCTTTTGAA TGCATGCAGT AATCGTTTCG GAAAACCACA GTTGGGATGC   3360

CCGTGCAGTT AATCTGTGTA ACAAACCCGT GCGCAGTTTT TGTAATCAAT ACATCTGGTT   3420

CAAGCAACAT GTTCGTGTCA GcCCGCTGAG CGTTCGACAC ACACTTACCT GGAAAGGGAT   3480

GTAGTTCCTT AATGAGGAGC AAAATATCTT TCACGTCATT TGACGAAACC TTCTGCACAC   3540

AAAGCCCCAT ACTATTAATC TGTGTCGTCA GCGCGTGCAC GGATACGCGT CCATCACACA   3600

TATTGTCAGA GCAGAAAAGC AATTCGCTGT GGTGTGTTAG TAGATTGATA ACACATCGAT   3660

ACAAGGGATC AGAGAAACGC TCAAAGCGCA gCCGCGCTTG GACTGCCAAT GATTCTTTAA   3720

AATTAAAAAC AGCACACCCT TGTGGCTCAA GTCTTTGAAT GAGCGCTATT GCCTGcGGTA   3780

TTTTTTCTTG AAGGGCTGTG GGCATACTAC CACACATGTT CTGAAAGATC GCAGGAGATA   3840

TGGAAAAAAA ACCGTGATCA TCTAACATCT GGATAAACGC GCACGCCAAA TCGAGCACAA   3900

TCGCTTCGTG TTTTTGATAA AAAACTTGTT CACGCAATAC AGCTCGGATA TTGTCAACCT   3960

GCTTATCGGG CTGATTTTCC AGCAATTGCT GAAAgCGATC ACGTGCGCGC ATGCgctCAC   4020

GTCTATCACC GAGCGACAGG TAACAAGCCT TCCCAGTGCG ACGAGCGGAC GAAGGACGTA   4080

TTTCTAAAAG GGGATTGCGT TGCACGGCAC GGAGAACCTC GGTCTTCAAA TCCCCCCGAG   4140

AAAGCTGCAG TAAACAAAGC CCGTGCACCA ACCGCTGATT GAGAACTAAC CGCTGCTGTT   4200

GAACAAGCTG CTGCATATCA CCACGTGCCC TGCAAACCAA TCCCACGAAA GGAATACAGA   4260

AAAGGAGCGA AACGCTCACA GGTGCGTAGT TTCTAGGTTC TCCATTTCTA CTAGCGCACG   4320

CGACGCCGTT TGCACATACG CTCTCGCCTG AGAAGCCTGC TCACGGAAGC TAGCATTGTC   4380

GCTCCCACGG ATATCAGACC AATTGGTAAG CGTTGCGTAG GCACTTGTTT TCGACGCCCC   4440

GATCATACTC GCCAACATAG CGCCAATGCA ATGAAATGCC TTTAAGGTTG CATCAAGCAA   4500

GACAGTCGCC TCAGAGTCAA TTACCTTTAT CAGAGCTTCC AGAATACCCT TACTTTTCAT   4560

GGTAACTGTG TGCTCTTTA   4579

(2) INFORMATION FOR SEQ ID NO: 73:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1015 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

```
TTCCCCAAAA CGAAGCGTCC AATCTTTTAA tAATGTGCAA GTTCaATATT CaAGGTACGT      60

ATTGGGAACA GCGCAGAACT CTGTTTCATT TCATCCCAAA AGTATAACGA TGAGACAAAT     120

GCATCCCACA GCTTAGTTTG TGGAGATTGC GATGACTGGG TGGGAAAGTA AGGCGTTAGT     180

TTCTCATTTT TAATCACAGT GTTTATAGAA AGATCGAGAA ATTTATAGAT GGAAAAAGTT     240

ATTAGCGGAG AAAAACTGAT ATGACTTTTA TCCAGATCTT TAAAATTAAT CTCTAACGAG     300

GAAGACAATG TCCCCTGTAT TTTTATACGC CGCTTCCAAA AACGTACTGT CAACGGAAAA     360

TCATCATGAG ACAACGTGAG CTTTAATTTA CTTATAGAAA GACCATTATT TCCAGAAGGT     420

GCAGCCTTGC CAGATTCCCC CTTTAGAAGA TAAGAAAGAG AAAAATACTT CCAGCCGAGC     480

GACACTTCAT AAGAATCGCT CATACTTTTT CCAATATCGT ATACGTAGGT TTGTTTACAG     540

GTTATTTCGT AAGGCATTCG AAAATCTAAT TCAGCGCGTA CcTGCGGTTT ATCGAGAAAG     600

CGTGCAGAGA GCGTCGCAGA GATATACGGA AAAGAAAGAA ACGTCGCAAT GCTATACGCA     660

TATGGTCCGG GCGCAAAATG CACACTGACA CGTATCTGCT GTGTATATCC ATACAGTGAA     720

AAAGCAGCAT TTGCGGTGAT GCTATGATCA CGTACATGTA ATGAACTTTT CCCAGTAGGT     780

CGAGACGAGT CGTATAGCAC GGGGGTAATT GACCAACCGA GAAAACTTTC TCGAAAAAAG     840

GAATCTGCAG AAAAAGGATA CACTGCAAGA TTATTCGAAC TCTGCACTAA AGCGGAGTTC     900

TGAAATCCAT TTTGTTTTGC CAGGTGGTTG ACTATGAATC GGATATCGGT GCTGCAGAGT     960

GCACTGTATC CCGTTTGACA TGCGAATTAT ATCATTTACA CACTGAAACT GAGAG         1015
```

(2) INFORMATION FOR SEQ ID NO: 74:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 9974 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

```
AAAACAGATT TGTAATGTAC CATCTGCCCA TGGATATGGT ATCTGCGGCG TCCGCGCAAG        60

CCTACCCCCG CAGCCCCCTG ATTGAGCGCT CGTCCCCTAC AGTCTCACAC TTTTTGCCGA       120

GAATATCTTA ACCGTGCTCT GTCAGCTCAA TACTTTGTCT ACAAGGAGAC GCGCCTGCCG       180

TGAGAATCCA TCAACGCTCT GCTCCGTGCG TGCCTGTGCT TCTTTTTCTC TTCTTGCCGA       240

GTGCGCCGCT TTGTGCGCGG GGTAGCAAGG ACTGGACGCC GCCGCAACTG GGCGAGGTGA       300

TAGAGAGTAC CGAGCAGGAC CTTGCAGAGT TTGATGCCGG CCTTTTCCGT GCGGATCGCA       360

TCCTGGATCG CCATGACCTC TACCGCAAAA CCATGCACCA GCTGTTCTCC ACGCTCCTTG       420

AAGAACCTAA AAACCACGCT AAGCACCTGC AGCTCATCGA AACGTTAGAA AAGCTCGCCG       480

GTCCAGAGAG CAAAGAAATA CACGAGTTTC TCAATCGAcT GCGCAATTCT TCTACGTACG       540

CATGTACGCT GCCCGTTTCT TTCACCTCAT GGAGCGGGCG CGCATCCTCA TGGCTCGCCA       600

GGAATACCTG AAGGCCGCGC TCCTGTACCG AAGCGGCTAC GAGCTCTACT ACGATGAGTA       660

CCTTGCCGAC CCGTCAAGTC CGGGGAAAAA GGAgGTGCGT GCTCGCGTCG AGCAsgCAnA       720

TGcGCATGTT TCCCGCGCAA AGCCCCTCCT AGAAGCGGTC GCCGCTGCAC GGGCTCAGTA       780

TCAGAACACG CAGAAAAGGA CGTATGCTGC CAGCGCCCAT GAaGGCTGCG CGCGCGCGCG       840

AcGCGTACTC TGCCGCCCCC GTGCGCCTGC TGcACCgGTG CCCGCGcACC GAGTGCAGCG       900

TATCCTCACT CCTTAACGGT GGAGGCAGAA TTAAGGATTT TGCAGGACTT TTCTAAAACC       960

ACTGAGGAAA GCGCCGCGCT CACTTCCCTG GTCCaAgCGC TTGGAGCGCT TTTAAAGTTT      1020

TCTCGCGACA TAGAGCACAC CGGTGTTGTT TTTGAACAGC TATCCACACG CGCGCAGAAA      1080

AATAACGAGA CACAAGAGGC CTTCTTGGCC GTTGCACGCA AAATTACGCT CGGGCGCAGT      1140

AAACTTGAGT TCGAAGGTAT TCTCGGCGCG CTCCAGGCTC CTGCCTTTGA CGCTTTTGTA      1200

GATCTTTTTG AAGCAGGTCG CGCACATGTA GCGGCGCTCC ACGACCAGGC GCGCGCACAG      1260

TTTACGTTTG CACATCCTCC GCACTCAGGC AGAAACATTC CCGCACCCAC CGACACTGcA      1320

CTGGCAAGTG CAGGCGCATG GGCAGCAGTC GGTGCAGGAC CTGCAGGATC GCTCATTCCT      1380

GGCGCTCCTC TCAGTGCGGG AGTCGGCTCT CGCGGCGCGT GGGGAGCGTT GCCTGCGCCA      1440

GTAGAGCCGC TGCTCCGCCA GGCGGATGAC GCATTGGGTG CGCTTGCAAG ACTGTGGGCA      1500

GCGTGCGCCC CGCTCGGTGC CCAGCATGGC AGATTTCCCC GCGATTATGA GACCTTTGGC      1560

GCGCAGATTG TAGCGCTCAG TGCGCACGCC GAgCGTTGCG CGCCACAAAA CACGCGTACG      1620

ACTTTTACCA TGCACTGCTC GCCTTCCAGC GCGCCCCCAC CGTGCCTGTT TCGGCTGCAT      1680

TGCGGCGTCA GGACCTTTCC CAGAATGAAG CGTTCGCGCG GgATCTGAGC GAACTTGCAC      1740
```

ACCACCAGGA GTTTTTGCGT CGTGCCCTTG CAGAAACCGA GTCTCTTTCC CCGCCTGCAG   1800

ATACGGCAAG CaCCCcGTCA CCGGGGGGTG CAGGGGATAC TCCAGTGCCC AGCCAGGCTG   1860

ATAAGGGAGG GGCAAAACAG AGCGCTGCCC CTGATACTGC GCAAAAGGCA GTAGCCCAAA   1920

AAGCGGGTGC GTCGGAGGAG GCTGACGCGT CGTCTTCCCC CTCCGAAATG GCGGTGCGTG   1980

CGGCGCGTGC ACAGCTGCAT GCCATCCAAA GTGAGCTGTT GCGCcGCTTC ACGcgCtTCA   2040

AACGCAACCG CTATACCGCA CATATGGCGT TTCATCAGCA CTCAGGCGTT TCTGCGCTCG   2100

CTGAGTATGC GCAgnAGcTT aCCAGTGCCG AGGAAGCATT GCGCTTTGAC GCGAAAGACG   2160

AGCGCAGGGT ACGCGCGTTG AGCTTTGTGT CTGAAACGGG TCCTCAGCAG GTGAGTAAGG   2220

ATATGGAGGC ACTTGATCGG TTGCTTTCTT TTTTTTCTGG CGAAGAAGAG TTCCTGTCTG   2280

AGCGTGGCTA TGCGTATGGG CTGCAGTCCC TGCGTGATTT GCGCACTCAG TTTGAACAGT   2340

TCTCTGCACG CGTGCAGACA CTTTTTTTGG CAGCAGAACA ACGGGCTATT CACGAACGAC   2400

TGGCGCGTCA AGAAGCAGAG TACCGTTACC GACAGGCAGT GGAAGGTCTA GGTCAAGATG   2460

ACTTTGGCGG TGCCCGTAAG AATCTGGTGC TATCTCGAGA AAAGGCCGAT TTGGCGCTCT   2520

CGTTGCGGTA CGACACCGGc tACGCTACCG AAACTGACAC GCGATTGAGC ACGCTTGATT   2580

CCTCAATTAA CAGACGGGAA AATGAACTGG TTGTAAAGGA CGTGCGCGCG TATATCgCAC   2640

AGGCAAAAGA TAAaTATTAC AaGGGAGAGG TGCTCGATGC GGAGcGTTTG CTCATTCGTG   2700

CGAAAAATCG CTGGGCAGTT ACAAACGTCA CCGAGAATGG GGAAATTACA AATTGGCTTT   2760

CTGTCATTAG TACGGCGGTT GCGCTCAAAA TCGGGCGGGT AATTCCTGAC TTTGCACCTC   2820

TTTACCCGCA GATGAGTCAG TTGTTACACC ATGCAGAGCA GCTGTACTTG CACGCGGCAT   2880

ATTTGAACGC GTCGCAGCGC CAAGAGATGG AACGGTTACT CGCCACCTCG CGAGAGAATA   2940

TACACAAAGT ACTGCTTGTC TATCCGTTGA ACGAGCGCGC AGGGCAGCTG AGTCTGAGAA   3000

TAGACCAACT GCTCGATCCC CGCTCCTTCC GGCAGCAGTT TGCAAAAAAG CTCGATACCA   3060

TCAGAGGAAC GTACAAAACC GAATCAAAAA AGGCCTACAG TTTGCTCCTA GATTTGTACG   3120

CAATCGATGC ACGCTTCTCT GGTATCGAAA AGCTGAAGCA GGAAGTGGAA ATCTACCTGG   3180

GGGTTCGATT GCCGCCGCCA AACCCGCAGG CCATTGCACA ATCGTCGAAT TTTACGCTGG   3240

CTGCGCGTCG TATCTTTGAG CGTAGAGACG CGGCGCTCTA TCAGGTAGCA ATTCAGCAGT   3300

TAGACGAGGC GCTTAAGCTG AATCCTGATA ACGATGCGGc TGCGCAgCTG AAAGATCGTA   3360

TCCAGTCGCT CACCGGTGAC GGTGCGGTAA ACGTACTCAG TAGCGAAGAC GAAAAAGAGT   3420

ATCAGCGCGC cTTGCAGGAA CTCCAAAAAG GAAATAAGCT CGTCGCCTCC GCGGTGGTTG   3480

```
AGCAGCTGTT ACAGAAAGAT CGCAATAAGA AGTCGGCAAA GATTCAGCAG TTAAAAAAGA    3540

GGATTGACGC ACAATTATGA ATGCTCGTCT GTGCTTTTTT TCGCGTCTTA TCTTTTGCGT    3600

ACTTTCTATC tGTGcTTtGC CACTTGTTGC TCAGGAAGAT AAGCTCTACT GGGAAGATCC    3660

GTGGGCACTC AGCACTGAcG TGCCGCTTTC GTCAAAGTTG CGTATTCGCA CGATGTCGTT    3720

GCCGTCGTAT GGCAGGAAGT GACGCCAAAA AATGCTACCT CGGGAGAAAT ACGACTGTCT    3780

GCGTCTTTTT ACGATGGCAG TACGTGGCAT ACCGTGCGTA CATTTCTCC ACCCCTTTTG    3840

TACAACCACC GTTCTCCTTC TCTTGCCTCC GTTGCTGTTA ACAGAAAAAA TGAGATTTTT    3900

GTTGCTGCCG CTTTTGATGC ACACACCATC ACCGTCTTTA AAACTACGGA TTTTGGAAAA    3960

TCATTTACGC ATACTGTATT GCGTTCTCAG GGAAGCGATA TTGTCGCCCC CTATGTGAGT    4020

GTTGCTTCAG ATGACTCGCT GCTGCTGTTT GCCTCTCACG GTTCTGAGGA TCACTTTTCT    4080

ATCTTGCTTT GCCGATCCGA AGATGGGGAG CGTTGGACTC CcTTTCAGGA GTTTTTGTCT    4140

ACCGAATTTA GCCGCAGACT CTTTTTGCCT TCGCATGTTT CAACGCAGGC CCAAGAAATA    4200

GTGGTGTTTC AGGCACATCA CCAAGAGGGT GAGAGAGCAA GCTATCAGTT GTATTCAACC    4260

GTTAGCTTTG ACCAGGGCAA TACGTGGTCT GCgCCTGTGC CTGTTACACA ACCTGATGAG    4320

TATCACAATC AGCGGCCCTT TTTGGATCGT CTCTCAGATG ATCGTTTTGC AGTTACGTGG    4380

GAGCGCTCTG AACGTACGTC GACGCGATAC GAGATGTGCT ATGCCGAGCT CGATCGCTAT    4440

GGGAGAAAAA TCGGGACTAC gCTCCGCCTG GCAGAACCTT CTGACCGTCT CATCACTCCC    4500

AACTTTGTGC ATATCGACGG TACCACATTC TGTGTGTGGG CAGGAGAGTC AGCCGGGCTC    4560

AATACCATTT TTCTCGCGCA GAAAAAGGAA GGCGCGTGGA GTACTACTGC CGTACGTTCT    4620

AGTGAGGATG CCTTGCTGTT TcCGCATGCG GTGCGCGTTG ACAATCACCT TGAGGTTTTT    4680

TGGCAAGAGG GAGAAGGGGC GCGTGCACGT GTGATGCGTT TGCGTCCAGA TCAGAGTGTA    4740

CAGCCACCGA CCCTGATTGC AGAAAATTTT TCGCCAAACG CGGTAAGAAA GGGGACGCGC    4800

GCGCGGtACG CATTGTATTT CCTCGGGATT CGTCAGGCAT TGCAGGGTAT AACTACGCGT    4860

GGCAATGCGG CGTGCAGCCT GCTGCTCCTC CTGATTACGT TGCACACTTT cCGGACAAAC    4920

CTCAGATAGA ACTGGAGGCA ACGCAGGATG GCACGTGGTT TTTGGCCGTA ACGGTGTGGG    4980

ACTTCGCCGG CAATAAGTCA GCTCCCGCGT ACCTTTCATA CACGCGGGGT ACTACGCCTG    5040

tGCGCGTCCA CAATTGCAAA CTCCTCTACT GGAGAACACG CATGCGCTGA AGAGCAACAC    5100

GTTTACACTC AGTTGGAATC AACCCAGTAC TGATGCGCAA GGAAACGAGG AGCGCGATCA    5160

CACCAGCTTC CTTTGGAGCT TACAACAGGT GGCACCGCTT TCAGCACTAA CGTCCCTGCG    5220
```

```
TGTGGATACT GATGTACGAA CGTTCGAAGA ATTTCAGCAG CGCTGCGTGC GCGCCTTTCC    5280

TATACCTGTG GAmGTGCACG GCAcGCGCAg CAGGcAGTCG TCCGTATCGT TCACTAATAA    5340

GGAGAACGGC ATCTATCGCT TTAGcGTATA TGCCCTTGAT CGCTCTGGAA ACGTGAGCGA    5400

GCCCGCAGTT GTCTTTTTTG CCTTACGGCA TTTCGTACCC TACACCGCCA TTCGCTATGT    5460

GGATGTGAAA AAAGATCCTG CCGGTTCATT GCAGATGTCG ATTGTTGGTA ATGGGTTTCG    5520

TGCGCAAGGG ACAGTCAGTC AGGTATACAT CGATCGGGAT CGCAAAGCTC CATATGACTT    5580

GGTATTGCAT GCGCAGGAGT TCGCCGTTGG TTCAGACAAC CTTATTTCAG ACATACACAT    5640

CGATAATTTA AAAAAAGGTT CTTACCACGT GGGGGTATGG CACCCTGCTC GTGGGGTGCA    5700

TTTTGCAGAG TCAAGAGTGA CGGTTTCTGA AATGGGAACG GTAAAATTCG GCGCGTACGA    5760

CTATGAGCAT CAGGTGCGGT GGAGTATCCC ACACACTGGT GGATTGAGAG TGAATTTTGT    5820

TTCACTGTTC ATGCTGATAG CGCTTTTTCT TGCGGGTGTG GTGTTTGCAG CGTCACTTAC    5880

CAGGATAGGT GATATCGTCG GAGAAGCGTT TGTACTTAAA AAGCAAGTGG AAGCGCTCAT    5940

GATAGGAGAG CTTATGCCGT CAGAGAAGAG ACGAAAGGCT ATGGCACTGA AAACACACGG    6000

TGCAGGATTG CGGGTGAAGT TCATCCTGTT TGCACTTACG CTGGTTATAT CTGTCATTTT    6060

TATTGTGTCC GTGCCGCTTG GAGTGCGGTT TTCAAAAACA CAAAAAGATT TGCTGGCTAA    6120

AAATCTTTTT TCTCGGGTTC AAGTGTTGCT TGAAAGTCTT GTGGCGGCAG GAAAGGTATA    6180

CCTTCCAGCG AAGAATAAGC TTGAGCTTGG CTTTTTGCCC AATCAAACAA CGGCATTGCA    6240

CGAAGCGCGT TACGCGtTAT CACAGGAGAA AGTGAAGAGC CTCACGAAGA AGGTATCGAT    6300

TTTGTGTGGG CAACGAATTT TAGCGATATT GAAACGGTGC TCAATGAGCC CGAATATCGG    6360

CAAGGCAATT CTCGTTTTGT TGACAAAAGG ATTGCGCAGA TTTTGCCGGC AATGGAGGAT    6420

TTGAACAGAC AGGTTAAGAA AGATGCAGAA AAGATAGCAA AGGGTATTGC GGATCTGACG    6480

CAGGAGGCAG TTGCGCTTGC GTTGCGCACT GATCAGGGGT CAGTACGTCG CCGAGATGAT    6540

ATTCAGTCCA TTACGCGGCA AATGGATCAA AGGCTTTTGG AAATTTTTTC TACATTTTCA    6600

AACAACGCGG TGGGCTCCTA CCCTGAATAT CGGGTTGATA ATTTATCAAA GCGTCACAGC    6660

TCCTACCTTT TCTATAAGCC CATCCTGTAC CGCCAACGCG GACACGcGgA TAGTTTTGTG    6720

CACGGCGTTG TGTTTGTAGA AGTCTCTACG CAGGAATTGC TCGAGCACAT TGAGGGTTTA    6780

CAGCGCGATC TCATTAAAAT GGTATTTTAC GTTTCTTTAA TCGCACTCGC CTGTGGGGTC    6840

TTTGGCGCGT GGATTCTTGC CTCTATTATC ATCAAGCCTA TACGCAGGCT GGCAAGTCAT    6900

GTGGCGATGA TTCGCGACAC GGAAAAAAAG GAAGAACTTG AAGGAAAACT GATTGCCATC    .6960
```

AAAGGGCAGG ATGAAATCGC TCTCCTCGGA AGAACTATCA ACGATATGAC AGAAGGGTTG   7020

ATCAAGGCGG CGCTTGCCTC AAAGGATTTG ACGGTTGGAA AGGAAATTCA AAAGATGTTC   7080

ATCCCGCTTG ATACCAACAC TGAAGGGAGA AAGCTTACAT CTGGGTATAC GTGCGATGAT   7140

CACGTGGAGT TCTTTGGGTA TTACGAAGGC GCGCTCGGCG TTTCTGGGGA CTACTTTGAT   7200

TACATTAAGT TAGATGATCA GCATTATGCC ATCATAAAAT GCGACGTTGC AGGAAAGGGA   7260

GTTCcCGCAG CGCTTATCAT GGTTGAAGTG GCAACGCTCT TCCAGAACTT CTTTAAAGAT   7320

TGGAATATTC AAAGTCATGG TATCAACCTA AGCGACATTG TCTCTCGCAT TAATGATCTC   7380

ATTGAGGCGC GCGGGTTTAA AGGAAGATTT GCAGCCTTTA CCCTGTGTAT CTTTAATACA   7440

GTGTCCGGTA CGGTGCACTT TTGCAATGCa GGGGATAATA TAATTCATAT TTACGATGCG   7500

CAGmAAAGAA AAATGAAGCG TATTACGtTG CGCAAACTTC TGCTGCAGGG GTATTCCCGA   7560

GTTTTATGAT TGATATGAAA GGTGGGTTTG GTGTGGAAAC CCTCACCCTG CGTACAGGTG   7620

ATGTCCTGTT CCTCTATACT GATGGCATAG AAGAGGCGAA cGTCTTTTTA GAAACAAGCG   7680

GTTTGAACTG GTAcTGTGCC AGGAACAGGG ACTTGCGCAT GATGCGCCCC ATGAGACACA   7740

TACGGTAGGT CAGGCCGGAG AGGAGCTGGG AGCTGAGCGT GTCAGCAGCA TTATCGAATC   7800

AGTCTTTCTG AGGAAAGGTT TTTCCCTACA AAAGTGGCAT AACCCTGTCG AAGGCGAAAA   7860

GTTTGAATTT GATTTCTCCT CTTGTGAAGG AAATCTAGAC GAAGCGGTGC TCGCACTTGT   7920

GGCGGTGGAG CAGGTGTTCC GTATGTATAA GCACCCTCGG GCAACCAACC TTGATAAAAT   7980

CAGGGTGGAT AAAAAAGTGG ATATGTTTTT AGCACGGTAT TTTGTTCAGT ACCCTGAGTA   8040

CTGTGCGCGC AAAGAGGTAA ACAGCGAGTA CGAAGAGTAC CTGTATTATA CGTTCATTAA   8100

AGAAGACGAC CAATACGATG ATCTCACTAT CTTGGGAATA AGAAAGAGAT AGTGCCGCTG   8160

TTGTGCAGGT TATTGCATGG TGTGTGGGTT GTGACAAGGA GACGCAATGC AGATTATACC   8220

CATTGCGAGT GGAAAGGGTG GGGTTGGCAA GAGTTTGCTT GCGGCAAATT TGTCCATAGC   8280

GCTCGGTCAA GCGGGGAAGA AGGTAGTAGT AGCGGATTTA GATCTTGGCG CGTCGAATTT   8340

GCATCTGGCG CTTGGCCAAA AGGGAAATAA GCACGGAGTG GGAACATTCC TTATGGGTGC   8400

CTCTTCTTTT GAAGAGATTA TGGTGCCAAC TGGATATCCC AATGTATATC TTGTGCCAGG   8460

AGATTCTGAG ATACCTGGCT TTGCTGCATT GAAGGTTTCT CAGCGGCGGG CTCTAACAGT   8520

GGGTTTGTTA AAAACGCATG CTGATTATGT GGTGCTGGAT TTGGGGGCAG GCACTCATCT   8580

TGGAGTGCTT GAGTTTTTTC TCCTTTCTTC ACGAGGGATT ATCGTTACTG AGCCTGCAGT   8640

TTcTGCGGTT TTGAATGCCT ACCTTTTTCT AAAAAATGTG GTGTTCAAAA TGTTGTGCGC   8700

TGCCTTTAAG AAAGGGACTG GGGGAAGTAT TTTTTTAGAG AATCTCAAGT CTGATGCTGC   8760

GGCGGTACAG CGCATGTATG TGCCTAAGAT TCTTGCTGAG CTTGAGCGTG TGGATCAGCG   8820

GGGAGTTGCA GTACTTCTGG ATCGGATGCG GTCTTTTAGG CCGAGACTAG TCATGAACAT   8880

GATTGCAGAT CCGAAGGATG TGGATAAGGC GTTAAAGATT CGCCGCTCGT GTGAGCAGTA   8940

TCTGAATATT ACGCTTGAGT ACCTTGGGGT CATATACCAG GATACGCAGC AGAATGTCGC   9000

GCTCTCCTCT GGTCTTCCCA TTGTTGTGTA CAAACCGCAG TCACTGATTG CCCAGGCAGT   9060

GTACCGGATT GCCGATAAGA TTTTGCAGTC AGAGGGTGAG GAGGCGTCTT CCATTGAGGA   9120

TTATGAAGGG TTGGTGGAAC GAAGTTTTGC CTCTGCAGAA GCAGAAGCAG AAGTGGATTT   9180

CCAGTTTCGT ATGGACTATC TTGAGGATTT GATAAAAAGC AAAACAGTGT GTGTGGGAGA   9240

TCTTGCTGAG ATCATAAAAG CTCAGCAGTA TGAAATTGCT ACTCTGAGGA AGCAAAATCT   9300

GCTCCTCCAA AGGAAAATAA ATAAGACATT GCGCAATGCG TGAACTTCAT GAGGGTGGGG   9360

TATAACCCCT ATTTGTGGGG GTGTTTTTGG GAGAATACAG TTTACgCGGA GyGtGGTGAA   9420

TGGTGACAGG ATAAACGGAA AACGGTGGCG GGGTAGTGCG CGGTGCATTT CCTTACGCGG   9480

TGGAGGTTGT GTGATGTTGA GTATTGTCTA TCCGTCGTGG ATTCGTCCGG AAATAATTCC   9540

TTCTTTTCCC TATTTTCGCT GGTACGGCTT CATGTATGTG GTTGCATTCA GTATCGCGTA   9600

CATACTGTTT CGCTACCAGG TGCGGCGCGG TGAGCTTGAT AAATGGAGTC GGGTGAGCGA   9660

GCCTGTCACG CAGGATGACA TTATGAGTTT TTTTACGTGG ACGATTCTGG GCATTTTAAT   9720

AGGGGCGCGT GTTTTTTCCA CCATGGTGTA TGAGGTCGAT TTGCTGTATA TGCGCAAgCC   9780

ATGGCTGATT TTTTGGCCGT TTTCTTTGCA AACGGGTGAG TGGGTTGGAT TGCGAGGAAT   9840

GTCGTACCAC GGTGGGTTAA TTGGCGCGCT CGTGGGGGGT GGcTTGTGGA CTCAGTCGCA   9900

TGGGAGAAGC TTTCTTGCAT GGGCCGATGT CGCTGCAGCG TCAACTCCAC TTGGGTATAC   9960

TTTnGnAGAA TTGG   9974

(2) INFORMATION FOR SEQ ID NO: 75:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 5861 base pairs
            (B) TYPE: nucleic acid
            (C) STRANDEDNESS: double
            (D) TOPOLOGY: linear

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

AGGAAGCACT GGAGCACGTC CGnAAGCACC GTCTCGCCCA TGCGCGTACC ACGGCAACAT   60

```
AACTTTGAGA TTCAGTATCC CCGGTCTCCG TGCACTGTGC AGTAAAGTGA TGCGTATACG    120

CTCCCTTTGG GAAATCCCAC ACATGCAGCT GCGCGAACAG CACACACTTC CAGCGTGCAT    180

CAGAGTCGCG TTCAAGACGC CTTACCTGAG TCAAGCCAAA CACAAGTCTA CGGACATTCC    240

CCCCAAACTT TTTACGCGCG CGCGrAwTTA CCGCGTGGTC CGGCACCTTA CCTACCCCGT    300

ACAAACTATC AAGCCCGATG TTCACAAACC GACGCTTCAA AAGCCCCCCG aGCAAACGGG    360

TGGCCGTCAG ATTCTCGTGC GTCAGGGrCC GTCCTGTCTT CCCATAATCT AGAATACATA    420

TCGTTGTAGA TACCTGCCTG CGCACGGGAT GCTGTTTTGG ATCAGACGAA GGCGCCGCAA    480

CGCGACACAC CAAATTAGCA AGCGCAGGCT CCCCCGCTGC GCTGTTTTCT TCCCCCTCGG    540

GGGACGGTAT GGTAGACACG ACAGgTTCGG GAGATGArGG TCGTTCCTGC AACGCAGAAT    600

CCTGTGCACA AAGGGAAGGA GCATGTACCA AAGCGGCAGT GTCAAACGAA AGCGTCACCT    660

CCTGCGAAAC GCCGCTGAcG GAAAGGGGAA AAAGAGCACT CCCTGcGTAT CTGCGCACAG    720

CGTCACCTGC ATCACATCGC CATGTTCAGG ACCAATTTGG TAGCATAGCG TaAAACGCCC    780

ACGCTGTAGT GGCCACACGC TGTCGCGATA CTGCAACGTG GCACTAACCC CTACATGGGT    840

CGGCGGTGCA ACGCGCGGAG cAAGGcGTAT ACCTTGcAGC AGCGCACGGA CATACGCGTG    900

CAGcGCAcGG CTaCGCGCCG TGTCAGTCTC AGAAGAAGGG TCATATAACG CTTCAAGCTC    960

ATTGAAAGCA GCGTATGTGT CCCCCGTACA GAGCGCATCC GTGATGCGCA TTACCCTCTT   1020

CTGTGTCAAA AAACGCCTGC AATGCCGCGT GGTGCTCGGC AGGGATATGT CCAACAGTAC   1080

CCTGcACCCA CGCAGACCCC CGCCCAGGAG AAAAAGACGC CGGGGATTCA GACACACTCG   1140

CATCTGCCGC GCCCGCCTCA TTCCCCATTG CCGCAGCATG AGAAACATGC TCAGAAGCAT   1200

GGGCCACACC CGAAGAATTC AGGCCATCCG CATTCAGGGA AAAACTTGAG CGTGTCTGTG   1260

ACGTACAGGC GCCAAAAAGC ACACTCCACA CCCACACCGT ACAAAAAAAA CGGTCCATGT   1320

ATAATCTACA CCTCTTTATT CTGCAGCGCA CACCACAGCC GCGTGCTAAA GTACCGTCAC   1380

GGGCCtTCGT TACAGCCACC CTACGATATC CACCAAAGAC ACATCACGTC TTCTTTTCGT   1440

TAGGGACGTG CAGGACGACG CCGACACCAC CCATACATGA ACGCAAAGCA AAATGCCCGT   1500

CGCATCGCTT CCGCACGGcT GCGCGCCTTC CGCGCCTCCC CTACCGCTCC TGCCTATCCA   1560

GCACACTCCC TTTTCACTGC ACAGACAGCG TACCAGCGTG CACGCACTGT GTTCTTCTAT   1620

GCGCCCCTGC CCCTAGAAAT AGACCCCTAC GCCCTTGCAT ACACTGCAGA AAACGCAGGA   1680

AAGCACGTAG CTCTTCCTCG CGTATCGGGA AACGACTTGC ACTTTCACGC AgTCACGTAC   1740

GCGTGCACTA CCCGCCCCTC TGTTTCCTGC TTCACCACCT TGTGCCCTAG GACCAGGGGA   1800
```

ATTAGAGAAC CCGATGCACA CAGTCCACGC CTCTACCCCC CGCACCCTTC GCCCAATACT   1860

CCTGCACAAA GAACACTTGC CCTACCGCTT TTGATCGTAG TTCCCGCACT GGCATTCAGC   1920

ACAAATGGCG CACGCCTCGG CCGCGGCGGA GGACACTACG ATCGCTTCCT CGCCCGGATC   1980

GCCGCTACCA TACCAGCAGG GAGCTACTAC ACGCTCGGCC TCTGCTTTGA TTGCCAAATC   2040

ATGGCTGTCA TTCCTCAAGA AGCACACGAC CAATCCGTAC ACGCGGTGCT CACCGAAACT   2100

CGTCTCATTT CCTGTGCCAC GGCGCGTGCA CCAGCGCCAC CGTTCTCTTT ATAGTGCCTT   2160

ATTCCTCCAT TCTAATCACA CACGTGCATG CACCAAGAGG ACAGCGCCGT GCTATCTTCC   2220

CAGAAAGGAG GATGAAAACA CGTGAAAACC ATTCTCATAC TGGGTGCAGG AACCATGCAA   2280

GCCCCTGCAC TTCGCGCAGn ACGGGAGCTT GGGCTGTGGG TGTGCGCGGT AGATGGGAAT   2340

CCGCATGCAc CsTGCGCGGC ACTTGCAGAC GAGTTTACCC CAATCGATTT GGCCGATAGC   2400

GCCGCGCTCG TncGCTnCAm gcGCcGCAAT TcGCGCGCrC sGCGGCTTGG ATGCTGTGTT   2460

CACCGCGGCA ACAGACTTTT CCGTTTCCGT CGCTGCCGTC GCCGAGGCCT GTGCACTCCC   2520

CGGCCACCGA TTGGAGGCAA CCAAAAACGC TACGGATAAA ACGCGCATGc gTGCCTGCTT   2580

CACACGCGCC CGACTGCGCT GCCCCCGCTT CACGTTCCTT GAGCCTGACT CGTTCGCCTG   2640

GGACACACCG CCTGGGCATG CCCGACTGTG TTCCCACCTG CATAGCGCTG GACTCTCGTT   2700

TCCTCTCGTC GTAAAACCGA CAGACAACAT GGGAGCCCGC GGCTGCACGC TCGCGCAATG   2760

CAAGGATACC CTCATAAATG CCtGCGCCGT GGCGCGCCAG TTCTCTCGCA GCGGCCGGGT   2820

GATTATCGAG GAATTTATTG TCGGAAGAGA GTTTTCCCTG GAAGGgCTCA TATTCGACGG   2880

GACGTTGTAC GTCACCGCAC TTGCCGATCG CCACATCTGC TTTCCTCCCT CATTCGTAGA   2940

AATGGGACAC ACGCTCCCGG CAGCGCTCTG TACACAAGAc GCACAAGCGC TCATCGACAC   3000

CTTCCACAAC GGTGTGCGGG CACTCGGGCT CACCCATGGC GCCGTGAAAG GAGATCTCTT   3060

CCTGAGTACC CCCTCCCCGA CGAAAACTCC ATCCACTGCC GCCACACCCA ACCCTTCTGC   3120

CCCGTACACA CCCGAAGCAG TATTGGGAGA AATTGCCGCA CGcCTTTCAG GGGGCTTCAT   3180

GTCTGGCTGG ACGGTGCCGT ACGCTCTGGG TTTCGACGTC ACACGCGCTG CATTGCACGT   3240

GGCGCTTCAC GGTCCTTCAG CTGCCGCCTC GGCTGCCACC GCGTCTGTCG CCCCCCCTCC   3300

TACTGCGCTc ACCtGCtGCG CACACAGCTC ACCACTCTGT CTCCTCTTCC AGAAAAAAGC   3360

CCATACGCCA GCGCAGAACG CGCGTGGATT TCCATTCCTG GGGTAATACA CCGAATCTGG   3420

GGCCTTGCAG ACGCTCAACA GATCGCCTAC GTCAAAAACG TGTTCGTACG TATGCAGGAA   3480

GGAGCCgcgG TGCGCTTTCC TCGTAATAAT GTGGAAAAAT GTGGCAACGT GCTGAGTCAG   3540

```
GCCCCCACCC GTGCACAGnT ATCGCCGCAG CAGAAACCGC GTGTCGCTGC ATTGTACTCC    3600

GCCTTGTTCC TGCACACCCT GCAACAGACG CCTTTCTAGC AAGAAAACGC AGCGCAGAAT    3660

CAGCGGCCAG CCCAGCGCTC CAGGACGCTG ATTCTGAGTA CGCAGCGTCT GCATCACACC    3720

CCTTTGGGCA AGAGAGTATA CCGGACATCG TCTGCGATGC CTCAGGACGC TTCTTTACCT    3780

CTGAGGTTGC CTGTGCACCG CTCGTGCGCA CAGGACTCTT CCTTATCCCC GAGCCACTGG    3840

TGCGCGctGA cGCACGAGAC GTGCAGGGTC GCAGCATCCA TGCGCTGTGT ACCCTTGCAC    3900

TTAAGGTAGA GCCTGCGCTC GAACCTGCGC TGTGCTTTGC GCGTTCCCAA AACCTCGCAG    3960

AGTTATGGCG CGCACTTATT CGCGGTGGCA TTCAAGGATT ACTATACGCG TTTGACTCCT    4020

TTCAACTGTC CTGATGTTCA GTGCCAGAAA AAATAAAACG CGTGCGCAAA AGCTCTGCGG    4080

cCTCAGTCAT ACGCAGCCAC GTCAGTCCCT GTGCACAGAG AGTCTGCACG GACGCGGCCG    4140

TTTTTTGCTC AAAACCAGGC ACCACGCTAC TACCGTCCTT CAACAAACTC ACGCGTGAGA    4200

GGACCCCTGC GCGCTGCATA TCCTGCAGCG TACACACAAC ACAATGCGAG AGTGCTTCAC    4260

CTGCTACGAA AATACGCTCA TGTGCGCGGA GCATTTGGCA GCGCTTAACA AAAAGAGAAT    4320

CCGCAGTCCC TTGAGACGCT GGATACTCCG AGGACAGCAC ACTAAATTGT TCCACACAGG    4380

GGTTTTCTCC TTTAAAGAAA AACTGAGGAT GTCTTGTACG ATGCGCGCGT TGCCAAAAGC    4440

GCACCGCCTC CACGATAAGC GGGTGCACCG CCTGTCCCCA ACTGCCACGC ACACAATGCT    4500

CGGGCCAAAG GTATAGAGGC CCCTTTCCGG TATATGCACG GAATGCCAAG TACCCCGCTA    4560

CGGTCTGTAC ATAGCCGACA CGCACAGGcA CACACGCCCC AGAACGCAAA CGTTCGAACG    4620

AAACTGTATC AAAAGGACCG AGAGCATCTC CTGTCAGGGA GCGCCAAAAA CAGGGGTGCG    4680

CAACGTGCAT CCGCGGGTGC CGATCGCAAc TTACGTACAA TGCATCCACA TGCGCAGCGT    4740

GCaCGCGCAA GAACTCAGCA ACGCGCACAC AGTCCTGATC CGCGCCGGGA ACGAACAACG    4800

CACCGCGTGG ATCGCAAAAA TCATTTTGAA AATCAACCAA AAAAAAGGCT CTGCTCATAG    4860

GGAAGCGCGC GCAGGAACGC GCACGcgCGT GCGCgcAnTn ACCCCACCCC AGCAGCACAA    4920

GCTAGACCCG GAACACAGGC CCTATCTGCA CCGTCAGCGG ACACTCCAAA CAATACTTTT    4980

GCAAAAACTG TTCCCATTTC AAATCAGTGG AGCCTGCGCC GTTCTTACAC TGCTCGAGGT    5040

TCGCAGTAAT CGGGATACCC ACACCGCTTG CCACACCGAC AGACAGTCCC ACAACCCCCG    5100

TAAAAAGAA ATGAAAGTCA AGGTTCACAG GAACACCGAC TATTTTGTCC TTGGTAGACA    5160

TGATATCCAC ACCCGTAGAA GGCAGAAAGA ACGCCCGCTC TCCCATACGG AGTACCCACC    5220

CCAATAGGAA CTGCGCGCGG AACAGGAGCG TAGAAAGCCC CGCATCTAGC TGcGTGACGA    5280
```

AAGTAAACCC GTTCTCCGCA GAAACCCCCA CCGAAAGCCC CAGCGTGGGG GTGAAGGCCA   5340

GTATATCGGT GCGTTTCGTA nCTTGGTTCC GCCCTCAnCG TTCGCGCCcT TTCCCCACAC   5400

AAaGAcTCCC ACCTGTCCAA cTCGGGGAGA AACAAAAACt GCGCGGCGTT CGCGTCAGTG   5460

CAAgACCCCA CAACACCGCC AAAAgAGCGC ACCCCCCCCC gCCACCGAAC GGCGCAgCGG   5520

CACATCACCT CACCCTCACA CCAACCACTC ATACACTACA CTCGGAACGT CGGCCCTATC   5580

GTGAGCGAGA GCGGCAAGGT AAACTCCTTG AAATTAAAGT CACGAACACC AACGGCCGTG   5640

CTCGCAGCAA CCGCAACTCC GGCAAAGGAA GTGAGATAGT ACTGCACCTC TAAGTTCAAC   5700

GGTACGCTGT ACAGCAGCTT GCTATACCAC GCAGACGATT TCCCCTCAGA TGTCGCACAC   5760

GAATCACCGC AGATATTCAC CCCACTGGAA ACGATGGCCC GCAGCCCTCC CACACGCACC   5820

GCGTAGCCAA TTAACGCCTG TGCACGCACA AAGACATTGG T   5861

(2) INFORMATION FOR SEQ ID NO: 76:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3694 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:

CGAGTAGGAG ATACATACCG ACACTCAGGG TTTACACACG CAGTATATGT GCCGACTCTG   60

CGATTTGATT TTTCAACCAA AAAGCATCGA CACTGAGGAC ACACTGCAAA GGTCGGTTTA   120

AAGTGAGTGA CAAAATCGCA GACAGGGAAA CGCGTGCAGC CATAGAATTC CTTTCTTCCC   180

CTTGTTTTTT TCCCCACGAT ATTCCCATCG CACGCAGGAC GCGGACACTT TGCAAGAGGG   240

ACAGGCTGAG TATTTCTACA CTCAGGAAAC TTTCCGCATG CAAGGAAAAA TCCAAACCTG   300

CCCAGTTTTT TCACCATCGT ATCACCACAC TGACTACACA CCACATCTGT TTTCTCATCG   360

AACACACCGC GCATGCTGTT AAGATCTTTC ATCACCGTTG AAACCTTTTC GCTGAAAGCA   420

GGATAGAAAT CCGCAATGAC ACAATTCCAC TTGATTTTAT CTTCCTCCAC CTCATCGAGT   480

TTACTTTCCA TGCGCGCGGT AAAACTTACA TCAACAACAT CATGAAAATA GGTGGTGAGA   540

AGATCACTAA TGACCTTTCC CAATGGGGTC GGCATTAGCT GTTTTTGAAT ACGAGTTACA   600

TAATAGCGAT CCAGCAGTAC TGAAATAGTC GGTGCATACG TTGAAGGGCG CCCAATTCCC   660

TTTTCCTCCA ACATTTTTAC GATACTTGCA TCCGTGTACC GAACAGGACC tGCGTAAAGT   720

GCTGTACGGA CTGCACGTTA TGTAGTGCAA CTACCTCACC TTCCTTCGTA GGGGGAAGTA   780

592

```
CAGCTTTAGA GAGATCTTTG GGGGATAACA TTTTCAGTAC ACGGTAGAAT CCCTGTTCAA      840

TAACCTGCGT TTCAGTTGCA CTGAAAACCG CCGGGCCAGC GGTAATTTCA AACGTCAAAC      900

TGCGCACTCT TGCATCTGTC ATCTGACTTG CAACAAAACG CTCCCAAATC AACGTGTACA      960

GACGTATTTG ATCACGCGTA AGGTGCGCTT TAATCCGCTC AGGAGTGTGG GCAACATATG     1020

TTGGTCGAAT CGCCTCATGT GCGTCCTGAG ACTTTCCCTT TGCAGCGTAC CGATTGGGAG     1080

TACCCGGCAG TGCGTCAGAA AAATGCGTTG CTATCCACGC GCGCACTTCC TTTACAGCAG     1140

CTTCAGAAAC GCGCACCGAA TCTGTACGCA TATATGTAAT GAGCCCCACG CGGTGGGTAC     1200

CAAGAGATAC GCCTTCATAG AGCTGCTGCG CAACCTGCAT CGTTTTACGC GAGGTAAACC     1260

CGAGCCTATT GGCAGCGCAT TGCTGCAACG TAGAGGTAGT AAAGGGCTGC TTCGGTCGAA     1320

CATTTTTTTC AAAACTGCGT ATTTGAGAAA CTCGTGCCTC ACTCTGAGAA AAAAGACCGA     1380

TAGCGCTTGT AGCCTCCTGT TTGCTTTTGA ATACAGCCTT TTTCCCTTGA ATCAGTATCA     1440

GTAGTGCAGA AAATGACTTT TTATCCTTTT CAAACGTTCC TTCAACCGTC CAGTATTCTT     1500

CTGGAACAAA GCGCTTTACT TCAACTTCTC GTTCACAGAT AAGACGAAGT GCAACCGACT     1560

GCACACGTCC TGCAGACAAC CCGTTTTTCA CCTTATGCCA CAGGAGCGGA CATAGGTGGT     1620

ATCCTACCAA ACGGTCCAGT AcGCGCCGCG CCTTTTGTGC ATTGACCTTT GCGGTATCTA     1680

TTGGAACCGG ATGGCCAATT GCCGCCCTAA TCGCGTGCGG TGTAATTTCA TTAAACACGA     1740

TCCTTTTGAT CGGCGTATCA CAATACGCCT GGATAGACTG TGCAAGGTGG TACGCAATCG     1800

CCTCCCCCTC TCGGTCACGA TCGCTGGCAA GAAACACTTG CAGTGACTGC TTAGATAGGG     1860

TGCGCAACTC TTTTAAACAC TGCGCACGAC CACGAACTGT AATGTACTCA GGCTGGAAAT     1920

CGTGCTCAAT ATCAATAGCT AAACGAGACT TTGGCAAGTC AATAACGTGG CCCATGGACG     1980

CTCGCACCAC GTAtGCGTTC CCAGATATTT TTCGATGGTC TGCGCCTTCG CAGGAGATTC     2040

CACAATAACC AAATGcTTCC GCGCAAATGT CTTCTGCCTT TTCGGTTGTA GCCCACGCAC     2100

TTCCATGTTT TCCGCCCCCT ATGCTTACCG AAACTGTCCT ACGTCCGACC CGTACATGCG     2160

TATTCCCAAT CGTCAAACAT ATCCTGcaCG CACGTCAGCG CGCGTGCACC TTCTGCATGC     2220

AGAAGTCTTC CTCCCTCATT TTGAAGACTT TCAAGCAAGG GTTCATACAC GTACACATCA     2280

CGTCCCTGTT CCAAGGCACA CAATGCGGTA ATCAACGCGC CTGACTTCTT TGGTGCTTCC     2340

ATAACAACGA GCGATCGTGC CAAACCTGAA ATGAGCCTAT TGCGTTCAGG AAAACGATAG     2400

CGCATCACAT GCTCAGATGG CGCATATTCA CTCAGAATGC ATCCTCCGGT TTCTATAATC     2460

CGTGCAGCAA GCGCGCTATT TGAGCGTGGA TATAACTGGT CTACACCACA GGCAAGTACT     2520
```

```
GCGAGGGTGT ACCCACCACC TGCTAATGCT CCTTTGTGAC AGAATCCGTC TATTCCACGT        2580

GCAAGTCCTG AAACAATGGC AATGCCCGAT TCAGCACACG CCTTGGAAAA GGCCAAACTG        2640

TTCCGAACTC CTTCACCGGT TGGAGTACGT GTGCCAACCA TCCCCACAAT AGGTTGGGTT        2700

GCACACGGCA AAGTGCCCCG ATAGAACAGC ACAAACGGTA CATCACTTAT TTCTCTTAAC        2760

CAGGAGGgAA ACGCATCATC GTCTTTGAAC ACCATTTTTA TCTGATAACA CCGCATGGTT        2820

TGTATACCGT GCCGAACGAG GTGAGGCAAC GCAGAAAGTT GCGCACCCGC AGTGcGTATG        2880

TGCCTTTCTA CAACACGCTC AAAATCACGC ACCTTCCATG CAGTAAGCTC CTGArAAGAA        2940

CCTACAGCTT TTGAAACGCG CAACCGCTCC CCACCTTTTA AAAGTGACA GTAAGAGAGC         3000

GCAAGAGCAA TCTTGTGCGT TTCAGTGAGT ACAGTATCCG TGTTCATCCA CGGTCCCCAC        3060

GCGTAcACGC TGCGATACAT TCGTGACTAA TTCTTTATTC TGTCGCGCAC GATCTCTCTG        3120

ATGGGAGTCA CTGCTTTGGA CAAGGATGCG ATCGTAACAC TCTATTGCCT GAGCATATCG        3180

TCCAAGTGTG TAGTATGCGC GCGCAAGAAC GAACAAAGCT GCTTCTGTAT TTTCTTTCAA        3240

TTCGAGGAGC TTTTGCAAAT GAGGAACCGC GTTCTGTGGC TGATTCAGTT CAAATACATA        3300

CAATACAGAA AGACCGTACA ACGCGTGGGA ATACTGCGCG TCCAACGAGA GCGCACGCAC        3360

ATACGCAGAG ACTgcTAACG CGCGATAAGC AGCTTGCTGC TGCATTTTCT TCTCAGGATC        3420

AaTAGGAGCG ACGTACTTAG CCGCATATGC GGCACACAAC GCCTGGTAAA AAAAAAGATG        3480

CTTATTTTCG GGAGCAAAGG TAATGGCCTG GGTAAACGCA TCAAGCGCGT GCGTATACAT        3540

CTTACGATCG AAGTAGCGCA ACGCGAGCAT CTTGTACCAA ACACCCACCT GATnCnCnGT        3600

GCGTGCAAAC GCTCGAGGCG CTGTTCGTGC AGCTTTACTG CCCTCCGCAG TTCTTCAATA        3660

GAGGTTGGAT GGGGCACTCC TTTCTCTAAA TCCT                                   3694
```

(2) INFORMATION FOR SEQ ID NO: 77:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 6422 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: double
          (D) TOPOLOGY: linear

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 77:

```
TTACCTaAAC CGCGAATTTC CATATGGTGA CCGaTACTTT TGTGCACCCC CCGCCCAATG          60

AGGCTATTTC CATTGACGCA AGAAATTTCT ACTAAGTCAT CTGCAACAAG TCGATGACCG         120

CGCTCAATTA ACTCCAGAGC AGTCTCACTT TTTCCTACTC CTGAATCTCC TGAAATAAGA         180
```

```
ATCCCGACGC CATACACCTC CACCAATACT CCATGAAGCG CTATCGTCGG TGCGAAGATA    240

TTGGAGAGAA CACGCATGAG ACGTAAAGAA AGCTCGCTCG ACGTAAGACG AGTGACCAAG    300

ATAGGGCAAG AAGAAGGCTC AGCAAGATGC AAAAACTTCT CCGGCGGGGT AATTCCATGG    360

GAAAAGATAC AACAAGGCAA GTCAAAGGTG AACATCTTTT CGATAGCACC GTATCGTCCC    420

TGCTCTAAAA GGGCGAGCAG ATACGCATGT TCTCCGCGGC CAAAAAGCTG GATCCGCCGG    480

TAGGnAAACA AGTCAAAAAA GCCTGACAGG ACAAGACCTG GTCGGTTCAG ATCCGAGATA    540

GTGATGGGAT TTGCCAGTCC ATGGTGACCT GCGATACAAC GCAGAtCAAG CGAATCGCGC    600

TCTTTCAGAT CGAGCTTGAG CACATCGAGA ACGGTAAAAA GAGGAGCACC CACGGCCGCT    660

ACTGTAGCAC AAAaGCCAGG ACCCGTAAAC GAGCCCACCG CAGTGAGAGC CTCTCTTCAA    720

GAGAGCCACA AGTCCGTACA GGAACTCTTG ACTTTTACAT ACAACTGGTG TCTGCTGACG    780

GCGCCCGGGT GTGGGCACA CGATCCATTA GCTCAGCGGA GAGAGCGGCC GCCTCCTAAG    840

CGGCAGGTCG GACGTTCAAG TCGTCCATGG ATCAGGAACG GCGATGGTCG GGCGAGGGGG    900

ATTTGAACCC CCGACCTCTC AGTCCCGAAC CGAGCGCGCT AcCACTGCGC TACCACCCGT    960

GCACGCAAAG AAACCACACA GAAAGGGACG CGCACCGCAC AGTGCGCAGA CGGGAGCGAC   1020

GGGGCTCGAA CCCGCGATCT CCGGCGTGAC AGGctGGCGC GATAACCAAC TTCGctACGC   1080

CCCCAGAACT TGCGCGCATC CTACATCACC CGCACAAAGT TATCAAGCGG CGATGATAGA   1140

TCACCCAAGG AAATAGCGGC AATAGGGATT GAACCTATGA CAGCGCGGAT ATGAGCCGCG   1200

TGCTCTACCA ACTGAGCTAT GCCGCCAAAA AACCCCCGAC CGCACACCAC CGCCATCCTA   1260

TCCCTTTTTT TTACACTCTG ACAAGTCCTG CCTTCCCCCT GCCTCCCCGT GCTTGACGCA   1320

AGAAACAATA AAATTGCTGC CTATGATTTC TTTAGCCGCG CGTATAATGC GCCGAGACAC   1380

CGGGCGTGGT CTTTCGTCGC ACAGTGGCAC TCGCGCTTCT TCTGCGTACG CTCCCCTGCG   1440

CCGCGCACTT CGGAAACCGT GACCGCACGT TCTACGACCT TAACAACGCG CCCCTTGCTC   1500

TGCGCGCCAT CCAGGACGCA TATCCTCATC TCAACGCGGT CATTGCCTAT GACCCGCGGG   1560

AACAGGACTG GCTCATCCGT TCAGACGGGC GCACCCTCTA CTGGGCAGAG GGGCGTCTTT   1620

TACCTCGAGA ACACCGTGAT CAAGCCCACG ACTGGCGCCC CATCATCGAT TATGTCTACG   1680

CGCGAGAAGT CCTAGACCCC GCGCACCTTT TTCCAGAAGA AATACACGCG CTTAGGCCTA   1740

AGACGCTTGC AATTAAACGC AGCGCTACAA AACCCTATCA CGACGCTTTT TTCACGTGGC   1800

TCTACGGTCC TGCCACACGT TCGGAAATCA ACGCTCGTCT CGCGCGCGAC TATACGTTCT   1860

TAGGAAAGCC CGTATACGTA CACAAAGCAC TCATCACACG CTTAAACGCA GTACAGGAAA   1920
```

```
AAATCCTCAC TGCCGCGAAG ACGCATGCTC ACGTACAAAA GTTTATCGAG GATCTTTTAC        1980

GCGTCGACGG CTTTAACTGG CGTGAGATTT CTGATTCTAG ACAAAAGAGT AACCACAGCT        2040

GGGGGATCGC GTTGGATCTT ATGCCCAAGA ATTGGCAACG CCACACCATG TACTGGAATT        2100

GGGAArctGC GCATAACGAA GATTGGATGC ACATCCCCAT AAAAAAGCGC TGGGCTCCAC        2160

CTGCAGAAAT CATCAGTCTT TTCGAAAGCG AAGGGTTTAT CTGGGGCGGA CACTGGATGC        2220

TGTGGGACAC TATGCACTTC GAATACCGGC CGGAATTACT CGCTGTACGT AAAATCCTTG        2280

CCGAGGGGAA CCGCTATGAC TTTCAAGAAC AAAATATAGT GGTGCATGCA GATGATTTTC        2340

CTGCGCAATA CTTTTCTCCC AAAGAAGTAT TCGGCACAGA TGAGAAGGAG CACATTACCT        2400

ATGCAGAATC CTGCGTtCGT GCAAcGCAnG CAmaGTGTTA AAGAACTCGT TCGTGCACGC        2460

ACGCTGGTAG CGCGGTTTTC TCCTATGCGT CGGCTGCACG TGTATGCACC TCCTGAAAGC        2520

ATTCACAACA GCATAGATAC AGCCCTATTA CGCATGACCG CACAACTGAA GAAAAATTAC        2580

ACAAATGCGA AAaTACGAA CAATTCTCGT TTGCTTTCAA AAAgCATGCT CAGACACGCG        2640

CGTCTCGCAG AAGCGCAGAT GTGTACACGG TATCGTGCCG TCATGCTCAG GCAATCGCAn        2700

TGCAGTATCC ACACGCCCTG TCTTGGCAAA GTAAAGAACG AAGCGATGCG CTGTGGATTG        2760

CGCTTTTTTC CgTACGGCAA GAAGCGGCAC GACGCTCCGT GTGCACACCC TCGTCTAAGG        2820

AACAATGCAT GACGCACGCA CTTTCTTCAT GCGTGGATCT TGCACGTACG CACATCCTGT        2880

TGCCATAGGG CGCTTCTTCC CCCTCTCTTC CCCTACTCAC ACACCACAGG GTACACTTAT        2940

GAAAAGTCAT GGCACCATGT GCTCAAGGAA TGCGCTTCTT TTGCCGAGAA GGGGCGCAGG        3000

GCTGCATGTT CTTACCCCAC GTATACgCGA GGCGCGACCG GTGAACACAG GCGTTAAGGT        3060

TATTCTCAGT CTATTCGCGA CGCTCGTCCT TATGGTGGGG GTGTTTTTCT GCGCACCACG        3120

CGCTTCTTTT GCCGAGTTTG AAAGACACTT TTACCAACCG ACTGTTCTCA GTGCGCTCTC        3180

TACCAACTTG CGTGAGGTCA GTAAGGCAAG TGAGGCTTGG CACAGTCGAT ATCGACCCCT        3240

GTTTTCTCAG TTCTGTGCGC TTGATGCAGT CAGAAGTAGT TTCGATCCTG CGCAAAAGGC        3300

TGAAGACATT ACACAACGTG CCCGGGAGGC CAGTGCGCTC TTGTCTTCTG TCGCTGGTCT        3360

CAAAGGGGTG CGTATTGTTG AGGCGCAGAA ACCAAATATC CATTTTTCCA CCTTTGAGTC        3420

CGACGTTCTC CTTGCTGACA GTGGTTCTGT AACCTACAGA AAGTACAACG CTGAGGAGCA        3480

CGACGTCCCT CTTCAGTTTC TAGGGGAGCA TTCCCCTGAA CCGAAGtTAT TATCGACGAG        3540

TACCATGATG CGCTGCTGTA CTCTTTCCCC TCCCTGGGGA ACTACGGGGA ATATCGTGGA        3600

CGCATTCTTT TCTACTTGTC CTTGCGTGCC TTGGGCACCC ACCTTATTGC GGAAAACAAA        3660
```

```
CTGAAGATCA CAGACAGCAT TGTTCCGCTT TCCGCTGATG ACtwaCCTTC GGTGGCATCG    3720

TTATTGGTAT CCCCCATGAG GGGGTACGTT CCCTCAAACC CTCTGTGCTC GCAGAGTGGA    3780

AGCGCAAGCA GTTCAGGGTA CAGACAGTCA GGAGTGAGCA GCACGAAGAC TGGGCACTGC    3840

TCAGTAATGC ATCAGGCGCC TTTGTCATTG CACAGGCAGT GCCCGTCTTG CTGTTTGGCT    3900

TTACCCCTCT GACGAAGGGC CTTGTCGCTA TGGTTGCTGT TGTGACTACT TTTTTGCTCG    3960

TATTCCAGTT GCTCAGCCTT CGCCAGGACC CCCTCACAAA ACTGAGGGAC AGGCTGATAC    4020

ACTTCCaCGC GCAGCTCCTA CACAGTTGTC TCGAACAGAA GGAATCACTC GAGTGGGAGG    4080

AGGTGCGAAC CCGACTTGAA CACCGCAGGC GGGAAACAGA TGCAGAAATG AAGAAGTCTC    4140

TTCCCAGGCG TCTCCGTATA AGGCGGGGAC GCGAGCTCGA TGCGCTCCTC AGTAAGGGTT    4200

GGGATGACGT CTTCTCCACC TTGGAGCATG GTTACGGTGG TGCGCGTGCT ATGAACCGCG    4260

CGCAAATCGA ACAGCTTGTC AGGGAAGTgc TCGCGCAGAG CCTTGCAAGT GGGGAGGCTG    4320

TGCTACCTGT GGCGATGCGT GCGGACACAG CCGATGAAGA GCTCGACGAG GTGCTAGAGG    4380

AACTCCCTGA CGAGGCAGCC TCTTTGCCTT CCGATTCCAG TCCGGAAGAG GACCTGGACC    4440

CCTTGGAGGA AGTCGAGAGT ATCGAGGGGA CTGCTGAAGA AAGCACACGC GAGTACGCGG    4500

CTGCGGGAGA CGCGCTCCTC TCGAAAACAC CCCAGCTTTC AACGCACAGC GAGTACGTGC    4560

CGGCGACACT CGCAGAACTC CTGGGCCGCA ACGCAGAGCC CGGCGACGTC GTGCGGGACT    4620

CAGCAGTCCT CGAATATATC GAAGGCtCTT CGACTATCGT CCCTGCTGTT TTTATGAGAG    4680

CCACGCTGTC CACGACTGCC TAGAAGTAGT CACGGGAGAA GACGGCCCCT CTCTCAGCCC    4740

TATGGAAAGC ATCGTCAGCA CCGAGGACGG TCTTTTCACC ATTCGGGTGA GTAAGGAGGA    4800

AGGAAACCAC CTCAACCGCG ATTTCAAGGC CCTGGTGGAT TCCGTACTGT ACTGAAGAAC    4860

ATATCTTTCC GCCGGTGGAG CCGGTCCTCT TACTCGGAAG CAGGCACGAA CGTGTGCGCC    4920

ACCACGTTGG TTTTTATGAG CTCATCGACT TCCGTCTGGG AAGACCTGAG CCAGTACTGG    4980

CTTCTGCCaA TCCCCCAGCT GATTTCCCCG CGCATCTGCA GCGTATCAAm CTTGTAGCTC    5040

CTCCCATCCG CcAGGGATGA AATTAATTTT GCAGTAATAG TACTTACCGC TTGCAGGATC    5100

TATTACGTAT CCACCACCCC AGGAGCCTGG AGAAGTACGC TCGAGGTTAT AGATGAAgGG    5160

CGTACCAACC AAGGGCATAT TGGCGACGTT TCCCTTCTTG GAGAAATCAG GATACGTTCT    5220

TGTACACGAA ACCACCGCCG CATTCGGAGC CCTCCCCATG CACACGAGGA TCTTGCCAAA    5280

GAGCTTACCA TCCTGAACAT ACAACCGCCA CACCCCAGTG GGTTTCCCTG TATTGTCATC    5340

AACGCTCTTC CAGATACCTT CCACCGGATC GGCCATTTCC TTCTGCACAG ACGGCACCTG    5400
```

```
CGCTGCCTTG TCCGAGCTTG CTGTGAAACA CGGCACACAC AGACACAGTA CGACACCATA    5460

TACGATAACC TTTCTCATAG CCCGCCCCCA CAAAATATAA TGCGCACAAC CAACAAGGGG    5520

AAATACGACC TGCAAAAAAG CAGGTACCAT ACGACGCACC CCCCACATAC ATCGAGCTCA    5580

CCGTGATTGT GCAAAACGCT CTTGAAACTC CAACACTACA TCCGATATCG GCGGAGCACC    5640

ATTAAGAGAA ACTAATTTCC CCCGCTCACT GTAAAAGTGG ACAATGGcTc CGCCTGCGCT    5700

CGATAGGCGG TAAGCCTCTG AAGAATTGCC GACATCTTGT CATCCTCCCG CACGACCAGT    5760

ACTCCTCTAC ACCGATCGCA CACACCCTCT CTCTTAnGsT GCGCAAAGAG CACATGATAA    5820

CTGCTCCCAC AGGCCGmACA CACCtGCGGC CAGTAAGACG CGCAACAAGG ACATCGTCCG    5880

GtACTACAAT ACTCACCGCG TAGTCTATCG GCACAATGTC CTCTAAGCAC CTAGCCTGCG    5940

TGACAGTGCG AGGAAACCCA TCTAGAATAA AACCGCTAAC CACATCTTCG TGACTGACAC    6000

GCTCCCGCAC TAGCTCCGTA ACGGTCTGGT CATCTACCAA GCCGCCCACT TCAACTACTT    6060

TTTGAACTTT TTTACCTAAT GCCGTCTGTT TCTGAATTGC TGCCCGAAGA ATACCCCCTG    6120

TGGAGATGTG CACAACGCCA CAACGCCCAG AAATTTCACC TGCAAGCGTA CCCTTACCGG    6180

CACCAGGAGG ACCAAGAAAA ACAAACCTCA TGAAACAAAC TCCACCTTAT CTTCCTACGG    6240

GGAAGAAAAA ACACACCTCA CCCCGTTCTC TCGCGCACAC AACCGACCCG AAGgCGTtAC    6300

ACGCCGCGCG CGCCGCGAAA CCCCGACCGC CCAAACAAGA GCGCACCGGT ACTCCAATTC    6360

TACACGGGGA GGATTCATTG GCAACACAAA ACTGCGACAT GCTCGATAnA nCCTTATGTA    6420

CA                                                                   6422
```

(2) INFORMATION FOR SEQ ID NO: 78:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 4646 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

```
CTTCAAAACC GCCAAACAGT ATAGAAACAA ACAGGTTTAT CATGTAAGGA AGAATTACGC      60

TCAGTAAAAC AAACGTAACT ACGATTGTAA GCGGAAGCCC TTCAGAGAGT AGATTCATTT     120

GTGGGGCCGC TTTTGTTAAT AAACCCATTG AAACGTGGAT TAACAGCAAT GCTCCCATGA     180

TAGGCAGTGC GATAGTCATC GCGTGTAAAA AAAGAGCACT CAACGCTTTG GTAAAAAACA     240

GCAGGAGCGC TTCCTGTTTC CGCAGAAAAA CAAAGCAATT AACAGCCTGA AAGCTCCGCA     300
```

```
GCACGCCTCC TAAAAACAGG ATTTGAAATC CTTTTATTTG CAAAAAAACA AGCATCGCCA   360

CAAAGTTCAA AAACTGTCCC ATCAAAGGAT TTTCTATTTG TGCAAAGGTA TCGTACATCT   420

CAGATGTTCC AAAACCCATC TGATACGAAA AAAACTGTCC TGCCGCACTA AAAGTCGTAA   480

AAATTACGCT AATAAAAAAA CCTGTTAAAA TCCCCAGCAA ACCTTCTCCG AGCAACAAAA   540

GCACATAGTA CGCACTAAAC TCACGAACCT GcATGGGTGC AGGGTACGCA AGCGGTAATA   600

CGAGGAATGC AATCAGGcCT GcGAGTGCCA CCCTCACTaC CCGAgaAACC GAgCGCACCG   660

AcAAGaGaGG TACCGTAAAC ATaAGcGcAA ACACGcGGAc CGCCGwCAAG aAAAAAAGA   720

GAAGCCTGAG aAAAgAGTGC ATCAAAGGAC CGTTCCATCG CACATATCCA CTAGACAGGT   780

CCACTCCTCA CTAACTGAGG GATAATGTCA AACAGCcTTA CGGTATAATT CTGCAGCATT   840

GTCAGCATCC ACCCACCGAG GAGGGCAATC ATTCCCAATA TGGTCAACAT CTTAGGAACA   900

AAGGTAAGTG TTTGTTCCTG AATAGACGTC ACTGCCTGAA AGATAGCCAC TATTAAGCCA   960

ACGACAAGCG CTGTGCACAG AACAGGCGCG ACAAGTAACA CCACCTGAAA AACACCCTCT   1020

CGTATCAAGC CTAATACCGC ACCTTGCGTC ATCACACACT CCCGTCCTGT ATGTTATAAA   1080

AACGAATGAA AAAGCCTATC TATCAGCAGA TTCCAACCGT CCACCAGCAC GAACAAAACC   1140

AATTTAAACG GCAATGAAAT CTGAACCGGC GGCAGCATAA TCATACCCAT AGACATCAAA   1200

ATACTCGCTA CAACCATATC AACAATTATG AAAGGTAAGT ACAGGAAGAT ACCAATCTGA   1260

AAGGCTACGG TCAGCTCATG CAGGATAAAA GCAGGGATAA GGACATACGT GGGCACGTCC   1320

GCAAGTGTAT CTGGCTTAGG CAGCTTTGCC ATGGACATAC AAAGACGCAC AGAAGACGGG   1380

TCATGCGCCA TCTGACGATA CATGAAGACA CGCAGCGGTC TTTCTGCCTC CGTATATGCA   1440

GTCTGGATAT CTACCTGGCC ATCGGTAAGA GGTTTAAACG ATTTGGCATA AATCTCAGTA   1500

AAGACCGGCC ACATGATAAA CAGGGCGAGA AACAATGCTA TGCCGTGTAA AACCTGTGTG   1560

GGCGGCACTT GCTGCAGCGA CAATGCACGT TTGATAAAAT CAAGGACGAT AGACAAGCGC   1620

AGAAAGGCAG TCATCAAAAG CAAGATACTC GGCGCAAGGG AAATGAGCGT GAGCAACAGG   1680

AGAAGTTGCA CAGAAAAAGC CACTTCCCGA TTGGTCTGGG GCTCCCGGAT ATCAAAATTG   1740

ATGAAAGGAA TGCGTGAAGC CGGCCGCTCA GCATTGATAC CAGTAACGCC GCGCTCGACA   1800

CCTCCGCGCG CATCCTGTGC AAAAAGCGGG AAGAAAAACA ATGACACGAA AAAGAGCGCG   1860

CGGCGTACGC ACGCACGAGC ACGGATCACA AAGCATCCTG TGCAGCAGAT TCTTCAGAAT   1920

CATTACGAGG GATGCGACGC AACTTCTTTC TCGTATCTGC AAGAAAATCT GCCTCAACTA   1980

ACGGCTTCCC CTTACCGGTA ACGCGCGCGG GCAAGAGGCG CGCGAGCATC TGAGAAAAAT   2040
```

```
CCGCACGGGC GTCAGTGCCC TGCTCATCAG CGACGATGTT CATGGTATCG ATGAGCTCTT    2100

TGTCCTTGAC CTCTGCAATG AGTGAAATGC ACGTATCAGA CGCTGCCAAT ACAAAGGCGC    2160

GCTCTGCAAG TCCTACCACG TACACTGCGC GCCCTGGCGc AATGGGCAAA CAGGCAAGCC    2220

GCTTCAAAAA TGGATCGTGC GCGCTAGAAA GAAACGcAtG CGTCTGATAA GACGCAGAAA    2280

CCCGTATACA GCCGCACACA CCACGCAGAG CACGAGACTA AAACGCAACA GAAGAGAAAA    2340

CACCGAAGGG GACGGGTCAC GCGCAACCGG GGCAGCGTCG AAACGGAACG CCTGTTCAGC    2400

AGGGGTGAGC GGGAACGCGT cCTCCCTCGT GTCACGCGCG GACTCAGGCG CCGGCTGCTC    2460

CTTCTGTGCA GAAGTCGCTG ACACCGCTTC TGAAACGGCA GAAACATCTA CCCCCTGCTG    2520

CTGTGCCCAG AGCTCAAAGG ACACATGCAA AAAAACATGC ACCGAAAGGA ACAGCGTcGC    2580

GCACCGsGGT ACGATCCGAA GGgAGcAATT AAACGTCCGC AATACGTTCT CCCGGCGAGA    2640

GAATTTCCGT AACACGCACC CCAAAGTTTT CATCAATAAC CACCACCTCT CCTTTTGCGA    2700

TCAACTTGTG ATTGACCAAA ATATCAACAG GTTCACCGGC AAGCTTATCC AACTCGATAA    2760

TGTGGCCTTC CCCCATACCC AGGATATCTT TAATCATCAT GCGTGTACGC CCGAGCTCAA    2820

CGGTAACTTC CATGAACACG TCCATGATAA GCCCGATATT TCCCTGTTCT GCGCCACCTG    2880

ctGCATTCTG CAGCGGATGA AACTGGACTG ACTGCACACT CGGACTCGCG GCGCCTATCC    2940

CCATCTGCAT GTTCACGCCC CCCATTTGAG AATTGCCCAC CTGgCtGCGG GCcTGCATTG    3000

CCCCCCCCAT CCTCTCGATA ATTCGAACCA TCAGCTGCTC AGACACCAAC TCCCACAGCG    3060

TATACGAAGT GCCATCTAGC TCCACCGTAT AGGTAAAAAC GCACAGACGC TGCGGGGGAA    3120

AGCGAACCAT CGCCTTAGGC ACCTGCACCG ACTCTGCAGG AGCCACACTT ACATTCTGTA    3180

CGTTCCGCGC CTCAAGCGTA GAAAGCTGTG CGCTGACATA TTGGGTGATC GTTTCACTAA    3240

CAACCGAAAG TCCCATATCA TCAATTTGAT CGTTGTCCTC ATGACTGACC AAATTGACGA    3300

GTTTCTGCGC AAACTCAGGA GCCATGAGGA ACAAATGGTC CCCTGnAAAn TCTCCTTCAA    3360

AATCGATGAC AgTTGCCACT AACATGTCCG GAATGACGCG GGAAAACTnT TCCTTAGAGG    3420

AAATTTCCAC ACGCGGCGGG GAAATAGAAA CAnTCTTACC GGTCAAAGAn TCCAAGCTCG    3480

GGCAAAAGGA ACCCACATTC GCCTGACAGA AAGACTGCAA CAACTCGCTT TGTGCGCTGG    3540

AGAGCCCCCC ACCGGAGAAA GACGCGCCCG CAGCGGGGGA GTCGCCGGCT CCCATCTCAA    3600

CACCTGAAAG CAGGGCATCG ATTTCAGCCT GAGAAATAGA GCCGTCACTC ATACAATTCC    3660

TCCTCGTCCG CGGATAATTC CTCAAAATCC TCTTGGGAGG TACTTTCTAT TCGTTCCAAA    3720

ATCTGCGCGG CAATCTTTTT TCCCACCACC CCAGGcTGrs AsrrAAACTT CTTGCGGTTC    3780
```

CCAATACTGA GCACAAAAGG ATCGCCCACA TGGGTGTCGT GCAACCGGAT GATATCCCCC 3840

ACCCGGAGCC CAAGGATATC GCGCACTGAA AGGCGGAGCG ACCCAACTTC TGCCACCACA 3900

TCCATATCCA CCGTGGATAG CTTGTCGCGC AGAACCCCCA TGTAtGCGTG GTAGAACTCC 3960

TGCGCACCGA AGAAAACCAA AACTGACTCG ACAACTTAGA AATGATAGGT TCTATGGTGA 4020

TGTACGGAAT GCAAAAGTTC ATCATCCCCT CTTCCTCACC TACCTTTGTC TCGAGCGTCA 4080

CCAACACCAC CATCTCTGAG GGAGGGACGA TCTGCGCGAA TTGCGGGTTC GTTTCAATTT 4140

GACCCAGGCG CGGACGCAGA TCGATAACcT GCGTCCAGGA TTCACGCACA TTCGCCAGAA 4200

TACGGACGAT GACCCCTTCC ATTACTGAAT TTTCAATATC AGTCAAATCC CGCTGCACCT 4260

TGGCTGCCTG TCCTGTTCCT CCAAAGAGGC GGTCAATGAT AGAAAAAGTA ATGGAGGGAT 4320

CCACCTCAAG CAtGCGTTCC CTTTGAGCGG ATCCATAGTG ATCACCGCAA GCGTAGAAGG 4380

CGTGGGAATA GAACGGATAA ACTCCTCGTA CGTGAGCTGA TCTACCGACG CAACGTGCAC 4440

GTGCACCATA CTGCGCAGTG cGCCGACAGC GAGGTAGTAG TCAACCGCGC AAAAGTCTCA 4500

TGCATCAACG ACAGTGTACG CATCTGCTCC TTTGAAAACT TATCTGGGCG CCTAAAATCA 4560

TAGAGCGTAA TCTTaCGGGT GTCGCTGATA GGGCGCGCAT CTTCAATACT TGmATCCCCA 4620

GAaCTGAtAG CCGTTAgCAG CTGAnT 4646

(2) INFORMATION FOR SEQ ID NO: 79:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 11191 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

ATGGAGTAAT GAGCAGTTTA CCCAGTATCT TGAATATCTT TTTCGGGTAC GCAGGCTGTC 60

TGCGCATACG GTTTCTGCCT ATGCGCGCGA CTTGAATCTT TTTGAACGCT GGTTGCAACA 120

CGCGCAGAGA GCGTGCGCGC GCGTAACAGT TTCTGATATG CGTCTGTTTG TGTGTGAGTT 180

AGGAAGACGG GGACTTTCCG CAGCGAGTAT TAACCGAGTT TTGTCCGTGG TGCGAGGTTT 240

TTATGTGTTT GCTAAAAAAA AACATTGGTG CGCGGACAAT CCTGCACGCT TAGTGAGGAA 300

TATAAAAGGT CCTTCAAAGT TGCCTCGTTT TATGTTTCCA CCGCAAGCAA AGGCGTTTTA 360

CACCTTACCA AGTCGTACGG ATATTTTGTG GCAGGAACGG GATGCGGCAC TTTTTGCGAT 420

GTTGTATTCA ACAGGATGTC GCGTTTCAGA GATAGCGGCG CTCTCATTGA AAGATGTGCA 480

TCCGCATCTT AGTTCTGCGA TTGTGCGGGG AAAGGGTGAT CGGGAsCGGA CCGTGTTTAT 540

TGCTCCGTTT GCGCAGAATT TTTTGCACGT GTATATGCAG GCGCGTGCGC AcGAnTGTGC 600

GCGcTACGCC TCTTGCACAC CCGCGCTGTT TGTGAATCAG CGGGGTCGGT CGCTTTCTGT 660

GCGCGGAATA CAGTACCTTG TTAGTCGGTA CGTGCTTTTG GCCCAGGACG TGCACGCGCT 720

GTCTCCCCAC GCGTTTCGGC ACAGTTTTGC TTCGACGTTG ATCCGTCGGG GGGCTGATGT 780

GCgCGTTGCg CAAGAGTTAT TAGGACATGC GAGTGTGTCT ACCACCCAGC GATATGTGCA 840

TGTGACTTCA GAGCAACTGC AGGACTTGTA TCACCGTGCG CATCCGCGTG GATAGGGGGT 900

AGGAACGGAG CGTCCAAACG ATGCGGGGAA GCGAGCTGCA GAGAATGTAC ACCAGTGCGA 960

AGTGCTTTTT TCTGAGACTT TTTTGAAGAA GACTTTCTTA AGCTCGCTTT TTTTGGTCG 1020

ACAATGGGTC GGGGGTAGTC GGATGAATAG TTTTACCAGA ACGGTGGATC TTTTGCATCG 1080

TGCTTTGGAT GTCAACgcGT TGCGCTATGA AGTGACGGCG AATAATCTTG CGAACGCAGA 1140

GGTTCCAGGG TTCAAGCGGA CGGACGTAAA CTTTGAAGCA GAGCTCAAGC GTGCTCTGGA 1200

TTCTCAAAGA AATGAGACAA GTTTTTTCAA GCAGGCAACT GCGGGGACGA ATATGTTGTC 1260

CAGTGATGTT ATCGACTAcC GcTCGGTGCG TCCGCGCCGC GTGTTAGACT ATTTGACGGA 1320

TGTGAAGGCG AACGGAAACA ATGTGGATGC TGAGCAAGAA GCCATGCATG TTCTCAAGAT 1380

TCAGATGCAC TATCAGATGT TGAGTCAGaT GGTAGGGTTC CAGTATCGTC AGGTTGAGTC 1440

CGTGTTACGT TAAGCGTATG GAGAAGCGTG ATGGGTTTGT TTAGTGGTAT CAATATTGCC 1500

GCGACGGGTA TGAGCGCGCA nTTTGGCGGG CCGATGTGAT CTCTGACAAC ATTGCTAATG 1560

CTTCCTCCAC GAGGACTCAA GAAGGTGGAG TGTTCGGAG GAGCAGGGTA GTTTTGGCGC 1620

AGAAGAATCC TGGCATTGAC TGGCGTATAC CTTTTGTGCC CGAGCAGTTG GATCGGGGGG 1680

TAGGCACAGG GGTTCGTGTG GTAAGCATAG AAAAGGACAA CGCTCCTTCT CGTCTTGTGT 1740

ACGACCCAAC GCACCCTGgA TGCGATTCTA TCAGGGCCGA AGtGGGgTAC GTGGAGTATC 1800

CtAACGTGGA TATTGTGACA GAGATGGTGG ATCTTATTTC TGCCTCTCGC GCGTATGAGG 1860

CAAACATATC AGTTATTTCA GGATCAAAAG AAAaTGTTTC AGCGTGCGTT GGAGATTGCG 1920

CGCTAGGTGT GTTGCGCGTA CAgTCTGTGA AGATGTCTGT GCTGTGTGAG GGGAGGATAC 1980

AATGACGCCA GTTGGTACCA TTACGAATAG TGCGAATGTA TATAAAGTTC CATCTCTGAG 2040

GAAGGTGCCT GAAATCGGTC CAGTGTCGGT AGAAAGCGTA AGGcAGCGCA TGCGAGGGAA 2100

TACTGACGCG GTGGATCAGG CAGTGAACAA AAAGGCGATG AGTTTTGAGC AAACGTTGCT 2160

GCGCGCTTTT GATCAGGTAA ATCAAAAGCA GCAGAAGACT GCTGAGTTGA CCGAGCAAAT 2220

```
GATAGTAGAT CCTGAGTCTG TTGACGTGCA TGATGTAACA GTGGCGATGG CGGAGGCTAG   2280

TATGTCCTTG AAAATCGCGC AGACTGTCAT TGATAAAGTC CTTAAGAGCT GGAACGATGT   2340

CACCACTGCT CGGTAAGGTT TACAAGGCCG GGCTGTTCTG CAAAAAGAGT ACCGACGGTA   2400

TATCAGgTGA AAAGAGGGTG GGACGCGCTT AGTGCGCATT GGCTCGTTCT ATAGTGAGGG   2460

GAGGGGACAC GCGTGGGCGA ATGGTTGGGG CAGCTCGGAG TCAAACTCAA AACACAGTGG   2520

AAGAAGTGGA CGCTCGTGCA GAAGTCTGTG CTTGCCGGCG CGGCGCTCGT GTCTGTCATG   2580

GGGGTTGTTG TCTTGCTCAC GTgGtCGcGA AGCCGACKcT CGTGCCACTT ATCGACACTC   2640

CTATCACTGA TGAGACGGTG CGGGAAAAGA TTATCCTGCG CCTTAACGAA GAGAATGTGC   2700

GTGCAACCGT CTCAAGCGTT GGGTTGATTT CTGTCTCGGA TGAGAAGACA GCGCGTCGTA   2760

TGCGCAGCAT CTTAATTCGC GAAGATTTGA TCCCAAAAAA TGTGGACCCA TGGGCCATAT   2820

TCGACGTCGA GCGATGGACG CGTACTGACT TTGAGCGCAG GGTGGACGTG CGGCGTGCAA   2880

TTAATAATAC CGTTACCAAT CATATCAAAG CGCTCGACGA CATCGATGAT GCCCATGTAG   2940

TAATAAACGT GCCTGAGGAT GCGCTTTTTC AGGCAGACCA GAAACCTATT ACTGCGAGCG   3000

TTGTCATTTT CCCTAAACCG TCGAGCACGA TCGCCTCAGA AAGAAAAAAA ATAGAAGGCA   3060

TTCAGAAACT ATTAAAGCTT GCAGTTCCTG GACTGAAGGA TGAAAACATC ACGATTGTAG   3120

ATAGTGATGC TACCGTCCTA AATGATTTTG AAGGGTTCAA GGACGCTGAT CGGCTGAGTC   3180

TCATTGAAAA GCAACAGAAA ATGATTGCGA arCTGgAATC CCAGTATGAG GCAAAAGTGC   3240

TGGCTCTCTT GCAAAAGACG TACGGTAAAG ACCGGGTGCG CGACTTAAAT ATCAAAATTG   3300

AAATGGATCT TTCTGAAAAG ACGTCGCAGA tACCAAGTAT CTGCCTATAG AAATCCGTCA   3360

GGACAATCCG GATACCCCGT GGGATGATTC TCAGGTTGTG CCCTCTGTCA CTTCGATATC   3420

TGAAACGGCm ACcACtACGT GGnCAGGGTA CGGGGCTTAA CCCTGAAGGA CCGCCGGGAG   3480

TTGAGGGTCA AACACCTCCT GCATACAAAG ACATGAGCAA CCAGGTGGGA CTTTCTAACC   3540

AGTCGGTCGT TAAGAAGCAA GAGGCGATTA GCAAGAGTGA GATCAACGAA GTAGTGAGCC   3600

CGGTGCTCGG CCGCAGGACG GTGTCGGTCA ATATCGATGG AGAATGGCGC AAAAAGAGAG   3660

ACGAGCACGG AAGATTCATT GTGAAGGAAG GACACATTGA ACGTGAGTAT ATCCCCATCT   3720

CTGnTGAGGA GCTGCGGGAG GCAACGAAGG CAGTGCAGGA TGCAATCGGC TTTGATGCGG   3780

GGCGTAAGGA TTCGGTAAGT GTTTTAAATA TCAAATTTGA CCGGACGTCA GAATTTGATA   3840

GAGAAGATGA GCATTACCTG CGCGTCCAGC AGAGGAACAT GATCaTCtTa TACTCCCtTg   3900

ccAGtgtgGC AATCGTTTTA TTTATCTTCA TGGTATACAA GGTTATCAGC AAAGAGGTGG   3960
```

603

```
AGCGTCGCCG TCGTCTGCGG GaAaGGAGCT TTTAAGGCAG CAGCAACTGA TGAGGGAGCG      4020

TGCCCTGTGG GAGGCTGAAC AGGCGGGGAT GAATGTTTCC ATGTCGGTGG AAGAGCGTAA      4080

GGnCTTGAAT TGCAAGAGAA TGTGTTGAAT ATGGCGCGGG AGCATCCGGA AGAGTTGCGT      4140

TGCTTGTGAG AACGTGGTTG ATGGAGGAGT AGTACTATGG CCGTTACATC CGTGAAGGAT      4200

AAGCTCGCCA CGGGAGAAAA AAAGCAACGG GATATCAAGT CTCTCAATGG TCGGCAAAAG      4260

GCAGCGATAT TTCTAGTTTC TATTGGGGAG GAAATATCCG CTAAGGTCAT GGGAGAACTT      4320

AAGGAAGACG AGATTGAAAA GTTGGTGTTT GAAATAGCGC GTACAGAGTC aGTTGATGCA      4380

GAACTCAAGG ATGCagTTTT AGAAGAaTTC CAGGAACtGA TGACCGCACA AAACTTTATC      4440

ACCTCAGGAG GTATCGATTA CGCGCGGGGA TtGTTGGAGA AGTCGTTGGG AAGTCAAAAA      4500

GCAATCGAGA TCATAAATCG GCTGACAAgC TCCTTGCAGG TGCGTCCCTT TGACTTTATT      4560

CGCAGAACTG ATCCCACACA CCTGTTAAAT TTTATTCAGC AAGAGCATCC GCAGACAATT      4620

GCGCTTATTT TGGCGTACCT TGAGCCGAAT AAAGCTTCTG TTATTTTGCA GAACCTCCCT      4680

GATGAGATTC AGAGTGATGT GGCTCGGCGC ATAGCCACGA TGGATCGGAC GTCCCCTGAT      4740

GTGTTGCGCG AGGTTGAACG AGTACTTGAG AAAAAATTGT CAACGCTTTC TAGCGAGGAT      4800

TATACGGCCG CAGGAGGTGT CCAGAACATC GTGGaCATCT TGAATTTGGT CGATCGTTCT      4860

TCTGAAAAAT CTATTGTTGA AGCATTGGAA GATGAAGATC CAGATCTTGC AGAGGAAATT      4920

AAAAAACGTA TGTTCGTGTT TGAGGATATT GTAATGCTCG ACGATCGGGC CATTCAAAAG      4980

GTGCTGCGGG AGGTGAATAT GGAAGAACTC GCAAAGGCAC TCAAGGTTGT CGACACTGAA      5040

GTACAAGATA AAATTTTTAG GAATATGTCG AAGCGGGCAG GGAGTATGCT GAAGGAAGAA      5100

ATGGAATACA TGGGGCCGAC CCGCTTGAAA GATGTGGAGG AAGCCCAGCA GAAGGTTGTT      5160

TCTATCATCA GACACCTTGA AGATAGTGGT GACATTGTCA TCGCGCGTTC AGAAGAAGAC      5220

GAGATGaTTG TGTAAATGTT GTTCCTGATA AGCGATATGG GGTTCGAAAG GAAGCAGACA      5280

GTATGCCAAA GmTsATATTT CGGAACCATG AAGTGAAGAA TCTTGATCAG TTCTTGCTGC      5340

TTGATCTGAG CAGGTCTTTT GGTGTCGAGC CTCAGATTGA GGAGGTGCAA AGCGAACCTG      5400

TGTGTCCAGT TCCTGATATG CGTGAAGTGC AAGAGGAAGT TGAGCTGTTT CGAAAAAGTT      5460

GGGAAGAAGA GCAGGTGCAG CTGCGCGCGC GTGCAGAGCG TGAGGCACAA GATCTAAAGG      5520

AGCGTGTAGA GGAGGAAATC ACAGCATATC GCGAACAGTG TACGCAGGAG GCGGATCGTA      5580

TCCTTGCTCA GGCAAAGGAA CAGTCTGAGC TACAAATTAG CGAGGCGCAA CAGCAAGCTG      5640

AACGCATGAT TGCTGAGGCA GAGACGTCTC GTCAGAAAAT ATGTGATCAC AGTAAGGCAG      5700
```

```
AAGGTATTCG TCTTGGCAAG GAAGAAGGGT TTCGTGCGGG ACAGGAAGAG GTGCGGTATT    5760

TAACTGAGCG TTTGCATAAG ATGATCGAAG AAGTGATGGG GCGGCGTCAG GGTATTTTGC    5820

GGGAAACCGA AAGACAGATT GTTGATCTGG TGTTGTTGAT GACAAGGAAG GTGGTCAAGG    5880

TCATTTCTGA AAACCAACGC GCTGTTATCA GCGCAAATGT GGTGCATGCG TTGCGTAAGG    5940

TGCGAACGCG CGGAGCGgTG ACGCTGcGGG TAAACCTTGC GGATGTGGAG CTTGTTACCC    6000

AGCACAAGCA GGAGTTTATC GCTGCAGTGG AGCGTGTGGA TGATCTAACG GTAGTGGAGG    6060

ACACGTCAGT GGGTAGGGGC GGTTGCgTGG TGGAAACGGA TTTTGGAGAG ATTGACGCGC    6120

GGGTTGCAAG TCAGCTCCAT GaGCTTGAGC AGCGTGTTTT GGAAGTTGCC CCCATTGTAG    6180

TGTCATCAAT GTCAGCATCT AAGGGTTCTT GATAGAGAAA GAGGCGTGGG TGTGCGTGTA    6240

TGGAAGCAGA CCTGTTGTGC AAGTATGAGG TGGCgCTCCG CGAGAGTGAG CCGGTAAAGT    6300

ACGTTGGGCA TGTGACAGCA GTGAGGGGTT TATTGATTGA AAGTCGTGGC CCTCACGCGG    6360

TAGTTGGTGA ATTGTGTCGG ATTGTGTTGC GCCGCCAGGG GCGACCGTTG ATAGCAGAGG    6420

TAGTAGGACT TGCaGGATCG ACGGTAAAAC TGATGAGCTA CACCGATACG CACGGGGTTG    6480

AAGTTGGCTG TGCGGTGGTA GCAGAAGGGG CGGCAtTTCA GTCCCCGTAG GAGATGCTTT    6540

ACTCGGAcGC GTTTTGAACG CGTTTGGGAA GGCAATTGAC GGGAAGGGGG AGATATATGC    6600

cgTCCTCCGC TCCGAGGTGT TGCGCGCGTC TTCTAATCCT ATGGAGCGTC TTCCGATTAC    6660

GCGTCAAATG GTAACAGGAG TGCGGGTGCT TGATTCkTtG CTGGCAGTTG GTTGCGGACA    6720

ACGTCTGGGT ATTTTTTCCG GTTCGGGGGT TGGGAAGTCG ACGCTGATGG GGATGATCGC    6780

GCGCAATACA GACGcAGATG TGTCGGTCAT TGCCCTTATC GGGGAGCGTG GCCGTGAAGT    6840

GATGGATTTT GTTGCGCATG ATTTGGGTCC TGAGGGTTTG AAGCGCTCGG TAATAGTTAG    6900

TGCGACGTCT GATGAAAnGT CCTTGCGCGG GTACGAGGTG CGTACACGGC GACAGCGATT    6960

GCAGAGTACT TTCGGGATCA AGGCAAACAG GTGCTGCTGC TGTTTGATTC TCTGACGCGC    7020

TTTGCAAAAG CTCAGCGTGA GATTGGGTTA GCGTCGGGGG AGCTCCCTGC AACGCGTGGA    7080

TATACCCCGG GGGTATTCGA AACGTTACCG AAACTGCTTG AGCGTGCAGG TTCTTTTTCC    7140

ATGGGGAGCG TCACCGCTTT TTATACTGTT TTGGTAGATG GGGACGATCT CGATGAGCCG    7200

ATATCAGACG CCGTGCGTGG AATTGTAGAC GGGCACATTG TACTCAGTCG CGCGCTTGCg    7260

cAGcgCAATC ACTATCCTGC AATAGACGTG TTGCAAAGCG TTTCTCGCTT GGCGCACCGC    7320

GTGCTGGGTG CAGACATGAA AGAGGCAGTG CGCATAGTGC GTCGTGCGCT TGCAGTGTAC    7380

GCAGAAGTAG AGGATTGGT ACGAGTTGGT GCGTACCAGC AGGGGAGTGA TGCAGAACTT    7440
```

605

```
GATCGAGCTA TTGCGATGCG CGCAGAGCTT GAACGGTTCC TAACGCAAGG AGCCCAGGAG   7500

CGCGTGCGTT TTCAGGATAC TGTAACGTCG CTGTCCATGC TGACAGGGCT CAGTATAGCA   7560

CAGCCGCCTT CGGGTGTGTG AATCTGCAAG AGCAGAGGAG ATAGCGCGTG TGAAAAGGTT   7620

TTGTTTTTCT CTTGAGCGTG TGCGACGCTT GAGAGCGTTT CGTGTACGCG AGCTGGAAGT   7680

TGAGTTAAGC AAAGTTCTTG CAGAATACGG AAGCATAGAT ACACAGATTC GATCGATTGC   7740

TGGCGAGTAT CGTGCGCGGA TGCAGGACGT AGCGCCAAAG CGTGGAGCAG TTTTTTCTGC   7800

TGCGTCGGTG AGCGCTGTGC AGGATCAAAT TGACGTGTTG CAATTACGCC GAGAACAGCT   7860

GCTCCATAAG CAGGCGCACC TTTCTTTTAC TCTTGAGCAA TTGCGAGAAC GATACGCGCA   7920

CGnGCGCCGT GCACACGAGG CTTTGCTCAT GCTTGAAGAA AAGGAGAAAA CACGCTGGCG   7980

AGAGCAGCGA CTGCGCGCTG AGGACCGAGC GTGTGACGAC CTGGTCAGCG CACGCGTaCC   8040

TGGTGCACCC AGCAAGCATT AATGGCTGGC GCGCTGCGTG CGCGCTcGGG TGTATGAGGA   8100

AGGCGGTCCA TGTCCGTGGA AGAGTATGAG CGTTTCGTGT GCCGTGCACG CTCGTTCCAA   8160

GATGGTGTCT GCCTCATTTC CCGCTTCTTC GTACCCTGCA GAACACAGAT CCCCCGTGAA   8220

CGCAAGGTGT GCAATACGGT ATAGGTAACA GCCATACGCA GGGGATGCAA AACAGGTAAC   8280

GGTAAAACCT GCGCAATTGA GGGAACAGTC TCCCTGTAAC AGGGTGCTCT CGTGGGCATT   8340

CCAGCACTTT TCCCCGCAAA AGATGTGCGG GGAGTATACG CGCAAAAGCG TGCGCACGCC   8400

CGCTTGGGTT GCGTCATCTG TGCGGGTAAG AAAAACAGCC GTCAGGGTGC AGGAATTCTT   8460

TTCAATGGTG TGGATCAAGT GAGTGCTCAC GTGCCCGGCA TCGATTAGGA GCGCTTCGTT   8520

TAAGGACAGG TTGCACAGCA GATAACTCGT GCGCTGGGTA TGCTGCTCTG CAGGGTGTAT   8580

GTAAACTTTC ATGGCAGGTC GCTCAGGTAT ATGCGGGCGT GTTCGGTACT AGATCCTTTG   8640

AATGTGCaGT GCGTAGTGTG CGAGGGCATA AAGTGCGTGT GTTGCACGGT GCAATTTTCA   8700

AAACGCACTG TTGCAAGCGT AGCGTGTAAG AAACGGGTGT GAGAAAGATC GCTCTCCTGA   8760

AAAGAAACAT CCTGCGCGTG GATACGGTTG AAGTTGTCAG ATAAGAGTAC TGCATGTGCG   8820

AAGCTCACCG TGTGCAGGAT GCTCCCCTCA AAAGAGGAGA ·ACTGAATGTC AGCTTCAGTG   8880

AAAGAGCTAT TGACCAGGTA GCAACGCTCA AAAAAACACA TACGCATCAC TACATTCGCG   8940

AACACACATG CATTAAATAC GGTATTAAAA AAATTGCAGC CGACGAAATG AAAACCCGAA   9000

AAGTCTACGC GGTTCAGGCG CATGCGTGGA GCGTACACGC CGTGTATGTG TGTTGACGTA   9060

TGCATGAGCT TAGCAAAGTC CGACGTAGAG CAATACGGTG TGGGTTCGAA CATGCGCGCA   9120

AAGCTTATAC GGGCTGGACG GGTGTGGTGT CAATGTATGC GGTGTGCAGG ACCTGGGGAT   9180
```